

For Reference

NOT TO BE TAKEN FROM THIS ROOM

Ex LIBRIS
UNIVERSITATIS
ALBERTAENSIS





Digitized by the Internet Archive
in 2020 with funding from
University of Alberta Libraries

<https://archive.org/details/Links1970>

THE UNIVERSITY OF ALBERTA

STOCHASTIC LINEAR PROGRAMS

BY



GRAHAM LINKS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF SCIENCE

DEPARTMENT OF COMPUTING SCIENCE

EDMONTON, ALBERTA

FALL, 1970

Thesis
1970 F
168

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled STOCHASTIC LINEAR PROGRAMS, submitted by Graham Links in partial fulfilment of the requirements for the degree of Master of Science.

Date . July 10, 1970

ABSTRACT

This thesis examines stochastic linear programs, i.e., linear programs in which some of the elements of the coefficient matrix, resource vector, or price vector are random variables. The three classical approaches to the stochastic linear program, linear programming under uncertainty, chance-constrained programming, and stochastic linear programming, are reviewed. The emphasis is on stochastic linear programming and the simulation of stochastic linear programs. A new estimator is proposed for the decision vector in a system of stochastic linear equations.

ACKNOWLEDGMENT

I wish to express my gratitude to Professor K.W. Smillie for introducing me to stochastic linear programming, to Professor U.M. von Maydell for encouraging my interest in simulation and statistics, and to Professor S.G. Ghurye for his invaluable suggestion regarding the first order estimator to a system of stochastic linear equations.

I would also like to thank the Department of Computing Science and the National Research Council for their financial assistance.

A very special thank you is due to my wife for her love.

CONTENTS

	Page
CHAPTER 1 AN INTRODUCTION TO STOCHASTIC LINEAR PROGRAMS	1
CHAPTER 2 LINEAR PROGRAMMING UNDER UNCERTAINTY	4
2.1 Introduction	4
2.2 Linear Certainty-Equivalence Theorem	8
2.3 Two-Stage Linear Programming Under Uncertainty	10
2.4 Example of Two-Stage Linear Programming	23
2.5 Inequalities for Linear Programming Under Uncertainty	26
CHAPTER 3 CHANCE-CONSTRAINED LINEAR PROGRAMMING	33
3.1 Formulation of the Model	33
3.2 Solution of the Chance-Constrained Model	36
3.3 Linear Decision Rules	40
3.4 Discussion	42
CHAPTER 4 SYSTEMS OF STOCHASTIC LINEAR EQUATIONS	44
4.1 Introduction	44
4.2 Transformation of Random Variables	47
4.3 Tintner's Approximation	53
4.4 Theorems on Matrix Algebra	55
4.5 A First Order Estimator	60
4.6 Simulation of Stochastic Linear Equations	70
4.7 A Necessary Condition for the First Order Estimator	84

	Page
CHAPTER 5 STOCHASTIC LINEAR PROGRAMS	88
5.1 Introduction	88
5.2 Analytic Solution	89
5.3 Simulation of Stochastic Programs	94
5.4 Preference Functionals	103
5.5 Active and Passive Stochastic Programming	109
CHAPTER 6 CONCLUSION	112
REFERENCES	115
APPENDIX 1 TESTING OF THE APL RANDOM NUMBER GENERATOR	121
A1.1 Introduction	121
A1.2 The APL Random Number Generator	125
A1.3 Description of Statistical Tests	128
APPENDIX 2 THE MONTE CARLO SIMULATION PROGRAM	163

LIST OF TABLES

TABLE		Page
2-1	Second-Stage Strategy for \underline{y}	26
4-1	Results for Basis #1 Using 1000 Random Equations	75
4.2	Results for Basis #1 Using 2000 Random Equations	76
4-3	Distribution of \underline{x}	77
4-4	Distribution of $\hat{\underline{x}}_1$	78
4-5	Results for Basis #4 Using 1000 Random Equations	79
4-6	Results for Basis #5 Using 1000 Random Equations	80
4-7	Discrete Density Functions for Sampling Technique	82
4-8	Comparison of Monte Carlo and Sampling	83
4-9	Results for Basis #2 Using 1000 Random Equations	87
5-1	Weighted Bases Technique	93
5-2	Summary of Simplex #1 Simulation	95
5-3	Generated Distribution of \underline{z}	96
5-4	Generated Marginal Distribution of x_1	97
5-5	Generated Marginal Distribution of x_2	98
5-6	Generated Marginal Distribution of x_3	99
5-7	Generated Marginal Distribution of x_4	101
5-8	An Alternative to Simplex #1	105

TABLE		Page
Al-1	Timing Estimates for the Function 'TEST'	127
Al-2	Frequency Table Produced by 'TEST' for 1000 Random Numbers	129
Al-3	Frequency Table Produced by 'TEST' for 50000 Random Numbers	130
Al-4	Frequency Table Produced by 'TEST' for 100000 Random Numbers	131
Al-5	Confidence Internvals for the Sample Mean in the Frequency Test	134
Al-6	Frequency Test Statistics	136
Al-7	Chi-Square Statistic for the Frequency Test	138
Al-8	Chi-Square Statistic for the Odd and Even Test	139
Al-9	Chi-Square Statistic for the Poker Test	143
Al-10	Chi-Square Statistic for the Gap Test	147
Al-11	Chi-Square Statistic for the Coupon Test	150
Al-12	Chi-Square Statistic for the Serial Test	153
A2-1	Data Set Specification for the Monte Carlo Routine	164

LIST OF FIGURES

FIGURE		Page
4-1	Average-Value Model for Basis #1	73
5-1	Average-Value Program for Simplex #1	100
5-2	Comparison of the Normal Density and the Generated Density for z	102
5-3	Generated Density Functions for z in Simplex #1 and Simplex #2	106
5-4	Generated Distribution Functions for z in Simplex #1 and Simplex #2	107

CHAPTER 1

AN INTRODUCTION TO STOCHASTIC LINEAR PROGRAMS

A stochastic linear program is a linear program in which some of the constants are random variables. Thus if a stochastic linear program is represented as

$$z^* = \max_{\underline{x}} \underline{c}' \underline{x} \quad (1.1-1)$$

subject to

$$A \underline{x} \leq \underline{b} \quad (1.1-2)$$

and

$$\underline{x} \geq \underline{0} \quad , \quad (1.1-3)$$

where A is an m by n matrix, and \underline{b} , \underline{c} , and \underline{x} are vectors in m -, n -, and n -dimensional real space (\mathbb{R}^m , \mathbb{R}^n , and \mathbb{R}^n), respectively, then some of the elements of A , \underline{b} , and \underline{c} may be random variables. In a stochastic linear program we must also assume that the joint distribution of A , \underline{b} , and \underline{c} is known, whereas in a classical linear program we assume that A , \underline{b} , and \underline{c} are known constants. Näslund (1965) is the only survey on stochastic linear programs, and while it provides a good introduction to the subject and contains a fairly extensive bibliography, it is lacking in detail. It is hoped that this thesis

will provide the missing details of the techniques involved and illustrate the techniques with simple examples.

If \underline{x} is chosen before the values of A , \underline{b} , and \underline{c} become known, then it is possible that \underline{x} may not satisfy (1.1-2) after A and \underline{b} become known, i.e., \underline{x} may be infeasible. There exist two methods which deal with this infeasibility problem - linear programming under uncertainty and chance-constrained programming. These two methods are discussed in Chapters 2 and 3. Chapter 2 also includes the details of using the decomposition algorithm in linear programming under uncertainty which was first suggested by Dantzig and Madansky (1960) but has not been elaborated elsewhere.

If \underline{x} is chosen after the values of the random variables are observed, then \underline{x} and z^* are also random variables. The study of the distributions of \underline{x} and z^* is referred to as stochastic linear programming. Stochastic linear programming is based on the solution of a system of stochastic linear equations, just as linear programming is based on the solution of linear equations. A first order estimator is proposed in Chapter 4 for the solution of a system of stochastic linear equations. This estimator is suggested as a possible

alternative to an estimator introduced by Tintner ((1955) p. 204). In Chapter 5 we will discuss stochastic linear programming and offer a suggestion for extending these estimators to stochastic linear programs. Further, since Monte Carlo simulations have been suggested throughout the literature as a method of analyzing stochastic linear programs, Monte Carlo simulations were used to analyse the examples in Chapters 4 and 5. As far as the author is aware, the majority of the material in Chapters 4 and 5 is original.

The emphasis in this thesis is on the simulation of stochastic linear equations and programs. It was originally hoped that these simulations could be done in APL. However, the long response time and space limitations proved to be too great a handicap. Therefore, after testing the APL random number generator in APL (see Appendix 1), it was translated into an assembly language subroutine (see Lewis, Goodman, and Miller (1969)) and then was used in conjunction with a FORTRAN program to simulate the stochastic programs (see Appendix 2). We also attempted to use APL notation as a substitute for the conventional mathematical notation in Appendix 1 by writing all the formulae in both mathematical and APL notation.

CHAPTER 2

LINEAR PROGRAMMING UNDER UNCERTAINTY

2.1 Introduction

Linear programming has proven to be extremely useful in a great variety of applications. However, the formulation of a linear programming model tends to be very restrictive and rigid. One reason for this is the deterministic nature of the classical general linear programming model. Usually a linear programming model may be written as follows:

Find

$$\max_{x_j} \sum_{j=1}^n c_j x_j \quad (2.1-1)$$

subject to the constraints

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i=1, 2, \dots, m, \quad (2.1-2)$$

and

$$x_j \geq 0, \quad j=1, 2, \dots, n. \quad (2.1-3)$$

In matrix form this model becomes

$$\max_{\underline{x}} \underline{c}' \underline{x} \quad (2.1-1')$$

subject to

$$A \underline{x} \leq \underline{b} \quad (2.1-2')$$

and

$$\underline{x} \geq \underline{0} \quad , \quad (2.1-3')$$

where \underline{c} is an n by 1 vector, \underline{b} is an m by 1 vector, A is an m by n matrix and \underline{x} is an n by 1 vector. The model defined by (2.1-1'), (2.1-2'), and (2.1-3') will be referred to as the deterministic linear programming model.

The assumption that \underline{c} , \underline{b} , and A are known constants is not always justified. In practical situations the values of these constants may not be known and they may be better represented as random variables. This lack of complete information is not enough in itself to make the deterministic model any less valid, as it still may be better to resolve the uncertainties by making some ad hoc specifications of \underline{c} , \underline{b} and A and then to solve the deterministic model. This method of replacing the stochastic elements by constants may or may not be justified depending on the sensitivity of the model to variations in the parameters. If this model is insensitive to parameter changes, then a deterministic approach may be justified, since the introduction of uncertainty into the linear model makes the analysis much more complicated. Stochastic models are inherently more complicated than the deterministic ones for at least three reasons. Firstly, if the stochastic model is to take into account all possible variations of

the problem, then the equivalent extensive form of the deterministic model will be large. Secondly, the objective function of a stochastic model is typically non-linear. Lastly, since the distribution of each random element must be specified, more information is necessary to specify a stochastic model. Since a stochastic model may be more realistic but more complicated than a deterministic model, it is necessary to make a trade off between reality and complexity.

Dantzig (1955), and Ferguson and Dantzig (1956) were the first to consider stochastic, linear programming models, and applied their model to the allocation of aircrafts to different routes subject to uncertain passenger demand. Their model may be stated as follows:

Find a vector \underline{x} to minimize

$$\min_{\underline{y}} E[\underline{c}' \underline{x} + \underline{f}' \underline{y}] \quad (2.1-4)$$

subject to

$$A \underline{x} + B \underline{y} = \underline{b} \quad (2.1-5)$$

$$\underline{x} \geq \underline{0}, \quad \underline{y} \geq \underline{0}, \quad (2.1-6)$$

where A and B are known m by n_1 and m by n_2 matrices, and \underline{c} and \underline{f} are known n_1 by 1 and n_2 by 1 vectors. First the n_1 -dimensional vector \underline{x} is determined, then the random m -dimensional vector \underline{b} is observed and lastly an n_2 -

dimensional vector \underline{y} is chosen to solve the equation (2.1-5). Dantzig (1955) also gives a computational procedure for this two-stage linear programming problem. The probability density function of \underline{b} will be assumed to be known. This problem could be referred to as linear programming under risk since the probability density functions are always known. However, the more usual name of linear programming under uncertainty will be used. In practice the derivation of the appropriate probability density may not be obvious and may have to be obtained from relative frequency data or subjective probabilities.

Dantzig and Madansky (1960) use the dual of the linear programming model under uncertainty to obtain a transformed model to which the decomposition algorithm may be applied. This paper and Dantzig (1955), referred to previously, are the main sources for the material presented in the remainder of this chapter.

Madansky (1960) has demonstrated that the substitution of the expected values for the random variables does not necessarily provide an optimal solution. However, such a formulation can provide a bound for the optimal solution. This bound and others considered by Madansky are presented in Section 2.4.

Managasarian and Rosen (1964) develop similar bounds to those of Madansky for the model containing stochastic, non-linear constraints. Wilson (1966) develops a generali-

zation of Dantzig and Madansky's method to obtain an approximate solution for the stochastic, non-linear model. These techniques involve non-linear programming methods and are not considered in this thesis.

2.2 Linear Certainty-Equivalence Theorem

Consider the following formulation of the deterministic, linear-programming model:

$$\min_{x_i} z = \min_{x_i} \sum_{j=1}^n c_j x_j \quad (2.2-1)$$

subject to the constraints

$$\sum_{j=1}^n a_{ij} x_j \geq b_i, \quad i=1, 2, \dots, m, \quad (2.2-2)$$

and

$$x_j \geq 0, \quad j=1, 2, \dots, n. \quad (2.2-3)$$

We will assume that the c_j 's are random variables with known distributions, and that the x_j 's must be chosen before the c_j 's are observed. As a result (2.2-1) cannot be optimized since z is a random variable. However, we may replace z by its expectation

$$\min_{x_i} E[z] = \min_{x_i} E \left[\sum_{j=1}^n c_j x_j \right]. \quad (2.2-4)$$

Other deterministic expressions could be used for z , e.g., the variance of z . Since the majority of papers consider

the expected value representation, it will be used exclusively throughout this thesis.

To illustrate the above type of problem consider the following example from Dantzig (1955). Suppose a dietitian wishes to formulate a minimum-cost diet. Since the prices may be subject to fluctuations due to changing supplies and demands, they are not known with certainty, and the deterministic model is not applicable. We shall assume that prices are random variables with known distributions. The dietitian must prepare a diet which not only meets specified nutritional requirements, but also has a minimum expected cost. Let x_j be the number of units of the j^{th} food purchased at a unit price of c_j , a_{ij} be the number of units of the i^{th} nutrient contained in a unit of the j^{th} food, and b_i be the minimum requirement for the i^{th} nutrient that is required by an individual. The problem can then be formulated as selecting the x_j 's to minimize the total expected cost (2.2-4) subject to the nutritional requirements (2.2-2) and (2.2-3). This problem can be solved by applying the Linear Certainty-Equivalence Theorem (Wagner (1969)):

Theorem 1 : Linear Certainty-Equivalence Theorem

Assume that a_{ij} and b_j in the formulation (2.2-2) and (2.2-3) are known, and c_j are random variables

independent of x_j . If x_j , $j=1, 2, \dots, n$, must be set prior to knowing the exact values of c_j , then x_j^* is a solution to

$$\min_{x_j} E \left[\sum_{j=1}^n c_j x_j \right] \quad (2.2-5)$$

subject to (2.2-2) and (2.2-3) if and only if x_j^* is a solution to

$$\min_{x_j} \sum_{j=1}^n E [c_j] x_j \quad (2.2-6)$$

subject to (2.2-2) and (2.2-3).

To solve the stochastic diet problem, the expected values of the prices need only be substituted for the prices and any linear programming algorithm applied to the resulting problem. Although the deterministic equivalent form of the diet problem is simple and easily obtained, it should not be assumed that this is always so. In general, the expected value may not be substituted for the random variables and the deterministic equivalents to the stochastic problems are not usually as simple as the diet example.

2.3 Two-Stage Linear Programming Under Uncertainty

In the two-stage linear programming under uncertainty we find an n_1 -dimensional vector \underline{x} and an n_2 -dimensional vector \underline{y} such that we minimize with respect to \underline{x}

$$E \underset{\underline{y}}{\min} (\underline{c}' \underline{x} + \underline{f}' \underline{y}) , \quad (2.3-1)$$

where \underline{c} and \underline{f} are known n_1 - and n_2 -dimensional vectors, respectively, and satisfy the constraints

$$A \underline{x} + B \underline{y} = \underline{b} \quad (2.3-2)$$

$$\underline{x} \geq 0 , \quad \underline{y} \geq 0 . \quad (2.3-3)$$

A and B are known m by n_1 and m by n_2 matrices, and \underline{b} is an m -dimensional random vector with a known distribution function independent of \underline{x} . In the first stage we select the values of \underline{x} knowing only the values of A , B , \underline{c} , and \underline{f} . After selecting \underline{x} , we observe a random vector \underline{b} . In the second stage we choose \underline{y} to satisfy (2.3-2). The vectors \underline{x} and \underline{y} are selected to minimize the objective function. In general, the solution will have a value of \underline{x} , which will be used regardless of the \underline{b} observed, but a complete strategy or contingency plan is needed for \underline{y} , i.e., for every possible \underline{b} observed there is a particular \underline{y} which will be used in order to minimize the objective function. It is apparent that the solution to this type of problem is more complicated than the deterministic one.

In the above formulation it is necessary to add the assumption that there always exists a feasible \underline{y} in the second stage, given a value of \underline{x} and any particular value

of \underline{b} , i.e., given any feasible \underline{x} and a possible \underline{b} , the set of linear equations

$$\underline{B} \underline{y} = \underline{b} - A \underline{x} \quad (2.3-4)$$

must always have a non-negative solution. Equivalently, we may assume that there exists a convex set K such that if $\underline{x} \in K$ then $\underline{x} \geq \underline{0}$ and there exists a \underline{y} for each \underline{b} such that $(\underline{x}, \underline{y})$ is feasible, i.e., $(\underline{x}, \underline{y})$ satisfies (2.3-2) and (2.3-3). Therefore, \underline{x} is said to be feasible if \underline{x} is in K , and (2.3-1) is equivalent to

$$\min_{\underline{x} \in K} (\underline{c}' \underline{x} + E [\min_{\underline{y}} \underline{f}' \underline{y}]) . \quad (2.3-5)$$

The following computational procedure, based on a method given by Dantzig and Madansky (1960) involves transforming the primal problem (2.3-1), (2.3-2), and (2.3-3) into the corresponding dual problem, and then applying the decomposition algorithm. If \underline{b} has the discrete probability density function $p_r(\underline{b} = \underline{b}_i) = p_i$, $i=1, 2, \dots, N$, then the extensive form of the two-stage linear program under uncertainty is

$$\begin{aligned} \min_{\underline{x}} & (\underline{c}' \underline{x} + E [\min_{\underline{y}} \underline{f}' \underline{y}]) \\ & = \min_{\underline{x}, \underline{y}_1, \dots, \underline{y}_N} (\underline{c}' \underline{x} + p_1 \underline{f}' \underline{y}_1 + \dots + p_N \underline{f}' \underline{y}_N) \end{aligned} \quad (2.3-6)$$

subject to the constraints

$$\begin{array}{rcl}
 A \underline{x} + B \underline{y}_1 & = & \underline{b}_1 \\
 A \underline{x} + B \underline{y}_2 & = & \underline{b}_2 \\
 \vdots & & \vdots \\
 A \underline{x} + B \underline{y}_N & = & \underline{b}_N
 \end{array} \quad (2.3-7)$$

and

$$\underline{x} \geq 0, \quad \underline{y}_1 \geq 0, \dots, \quad \underline{y}_N \geq 0. \quad (2.3-8)$$

Conditions (2.3-6), (2.3-7), and (2.3-8) define a deterministic linear program which can be solved using the simplex method. However, if N is large the problem can soon get out of hand since there are Nm constraints and $n_1 + Nn_2$ variables. For this reason Dantzig and Madansky (1960) suggest applying the decomposition method to the dual problem.

The dual to the extensive form is

$$\max_{\underline{n}_i} (\underline{b}'_1 \underline{n}_1 + \underline{b}'_2 \underline{n}_2 + \dots + \underline{b}'_N \underline{n}_N) \quad (2.3-9)$$

subject to

$$\begin{array}{rcl}
 A' \underline{n}_1 + A' \underline{n}_2 + \dots + A' \underline{n}_N & \leq & \underline{c} \\
 B' \underline{n}_1 & \leq & p_1 f \\
 & \vdots & \vdots \\
 B' \underline{n}_N & \leq & p_N f
 \end{array} \quad (2.3-10)$$

The vectors \underline{n}_i are unrestricted in sign. If

$$\underline{n}_i = p_i \underline{\gamma}_i , \quad i=1, 2, \dots, N , \quad (2.3-11)$$

then the problem becomes

$$\max_{\underline{\gamma}_i} (p_1 \underline{b}_1' \underline{\gamma}_1 + p_2 \underline{b}_2' \underline{\gamma}_2 + \dots + p_N \underline{b}_N' \underline{\gamma}_N) \quad (2.3-12)$$

subject to

$$p_1 A' \underline{\gamma}_1 + p_2 A' \underline{\gamma}_2 + \dots + p_N A' \underline{\gamma}_N \leq \underline{c}$$

$$B' \underline{\gamma}_1 \leq \underline{f}$$

$$B' \underline{\gamma}_2 \leq \underline{f} \quad (2.3-13)$$

$$\vdots$$

$$B' \underline{\gamma}_N \leq \underline{f} .$$

Except for $\underline{\gamma}_i$ being unrestricted in sign, the problem is now in the form required for the application of the decomposition algorithm. The first constraint represents the major problem, while the last N constraints can be used as a single subproblem as Dantzig and Madansky (1960) suggested.

However, in the model (2.3-9) and (2.3-10) the first constraint can be used as the major problem, and each of the N constraints

$$B' \underline{n}_i \leq p_i \underline{f} \quad (2.3-14)$$

can be considered as a subproblem. This approach may be extended readily to include problems where B is random. If B_i is the random matrix corresponding to b_i , then the i^{th} subproblem is

$$B_i' n_i \leq p_i f . \quad (2.3-15)$$

The following analysis is more general than that given by Dantzig and Madansky (1960).

In the original dual problem, the dual variable n_i are unrestricted. In order to apply the decomposition algorithm or the simplex algorithm the variables must be non-negative. If we let $n_i = \alpha_i - \beta_i$, where $\alpha_i \geq 0$ and $\beta_i \geq 0$ for $i=1,2,\dots,N$, then from (2.3-9) and (2.3-10) we obtain the following problem:

$$\max_{\alpha_i, \beta_i} (b_1' \alpha_1 - b_1' \beta_1 + \dots + b_N' \alpha_N - b_N' \beta_N) \quad (2.3-16)$$

subject to

$$A' \alpha_1 - A' \beta_1 + \dots + A' \alpha_N - A' \beta_N \leq c$$

$$\begin{aligned} B_1' \alpha_1 - B_1' \beta_1 &\leq p_1 f \\ &\vdots \\ B_N' \alpha_N - B_N' \beta_N &\leq p_N f \end{aligned} \quad (2.3-17)$$

and

$$\alpha_i \geq 0 , \quad \beta_i \geq 0 , \quad i=1,2,\dots,N . \quad (2.3-18)$$

In order to simplify this problem let us define the following vectors and matrices:

$$\underline{\pi}_i = \begin{bmatrix} \alpha_i \\ -\beta_i \end{bmatrix}, \quad (2.3-19)$$

$$\tilde{A}' = [A' \mid -A'], \quad (2.3-20)$$

$$\tilde{b}'_i = (b'_i, -b'_i), \quad (2.3-21)$$

and

$$\tilde{B}' = [B' \vdots -B']. \quad (2.3-22)$$

The problem becomes

$$\max_{\underline{\pi}_i} (\tilde{b}'_1 \underline{\pi}_1 + \tilde{b}'_2 \underline{\pi}_2 + \dots + \tilde{b}'_N \underline{\pi}_N) \quad (2.3-23)$$

subject to

$$\begin{aligned} \tilde{A}' \underline{\pi}_1 + \tilde{A}' \underline{\pi}_2 + \dots + \tilde{A}' \underline{\pi}_N &\leq \underline{c} \\ \tilde{B}' \underline{\pi}_1 &\leq p_1 \underline{f} \\ \tilde{B}' \underline{\pi}_2 &\leq p_2 \underline{f} \\ &\vdots \\ \tilde{B}' \underline{\pi}_N &\leq p_N \underline{f} \end{aligned} \quad (2.3-24)$$

and

$$\underline{\pi}_i \geq 0, \quad i=1, 2, \dots, N. \quad (2.3-25)$$

The first constraint represents the major problem and each of the remaining N constraints represents a subproblem. It

is necessary to assume that each subproblem is a bounded set; if not, certain modifications can be made to enable a solution to be found (Dantzig and Wolfe (1960)). The decomposition algorithm can now be applied to this model.

If S_i is the set of $\underline{\pi}_i$ such that

$$S_i = \{ \underline{\pi}_i | \tilde{B}' \underline{\pi}_i \leq p_i \text{ f} , \underline{\pi}_i \geq 0 \} , i=1,2,\dots,N, \quad (2.3-26)$$

then S_i is a bounded convex set, and there exists a set

$$W_i = \{ \underline{\pi}_{il}^*, \underline{\pi}_{i2}^*, \dots, \underline{\pi}_{ih_i}^* \} , i=1,2,\dots,N , \quad (2.3-27)$$

$$\text{and } h_i \in \{1, 2, \dots, n_2 + 2m\}$$

of extreme points of S_i . Since S_i is a bounded convex set, any point $\underline{\pi}_i$ in S_i can be expressed as a convex linear combination of the extreme points of S_i , i.e.,

$$\underline{\pi}_i = \sum_{j=1}^{h_i} \lambda_{ij}^* \underline{\pi}_{ij}^* , \quad i=1,2,\dots,N , \quad (2.3-28)$$

where

$$\sum_{j=1}^{h_i} \lambda_{ij}^* = 1 \quad \text{and} \quad \lambda_{ij}^* \geq 0 , \quad i=1,2,\dots,N , \quad (2.3-29)$$

$$j=1,2,\dots,h_i .$$

Also any point which is not in S_i cannot be expressed in such a form. Thus the two equations (2.3-28) and (2.3-29) are equivalent to the n_2 constraints

$$\tilde{B}' \underline{\pi}_i \leq p_i \text{ and } \underline{\pi}_i \geq 0, \quad i=1,2,\dots,N. \quad (2.3-30)$$

Moreover, the representation of $\underline{\pi}_i$ in terms of the extreme points of S_i can be substituted into the model (2.3-23), (2.3-24), and (2.3-25) to yield the model:

$$\max_{\lambda_{ij}} \left[\tilde{b}_1' \sum_{j=1}^{h_1} \lambda_{1j} \underline{\pi}_{1j}^* + \dots + \tilde{b}_N' \sum_{j=1}^{h_N} \lambda_{Nj} \underline{\pi}_{Nj}^* \right] \quad (2.3-31)$$

subject to

$$\tilde{A}' \sum_{j=1}^{h_1} \lambda_{1j} \underline{\pi}_{1j}^* + \dots + \tilde{A}' \sum_{j=1}^{h_N} \lambda_{Nj} \underline{\pi}_{Nj}^* \leq c, \quad (2.3-32)$$

$$\sum_{j=1}^{h_i} \lambda_{ij} = 1, \quad i=1,2,\dots,N, \quad \text{and} \quad (2.3-33)$$

$$\lambda_{ij} \geq 0, \quad j=1,2,\dots,h_i; \quad i=1,2,\dots,N. \quad (2.3-34)$$

This model may be represented as

$$\max_{\underline{\lambda}} \underline{c}^* \underline{\lambda} \quad (2.3-35)$$

subject to

$$A^* \underline{\lambda} \leq \underline{t}^* \quad (2.3-36)$$

and

$$\underline{\lambda} \geq 0. \quad (2.3-37)$$

The problem now has $n_1 + N$ constraints, (2.3-32) and (2.3-33), as compared to the Nm original constraints (2.3-7). However,

we must now solve for K variables (λ_{ij} and π_i) where

$$K = \prod_{i=1}^N h_i N_m . \quad (2.3-38)$$

If the revised simplex method is used, it will not be necessary to find all the vectors in each w_i since there are only $n_1 + N$ basic vectors in the basis at anyone time.

Let the element in the i^{th} row and j^{th} column of the matrix A be represented as $(A)_{ij}$ and let the i^{th} row and j^{th} column of A be $(A)_{i;}$ and $(A)_{;j}$, respectively. Finally, let the matrix formed from the first j columns of A be $(A)_{1;j}$.

The first step in the decomposition algorithm is to find the minimum element in the vector

$$(z_{jk} - \tilde{b}_j' \pi_{jk}^*) = \underline{c}_B B^{-1} A^* - \underline{c}^*, \quad (2.3-39)$$

where \underline{c}^* is the vector of coefficients of the objective function (2.3-31) or price vector, and \underline{c}_B is the vector of prices corresponding to the basic variables. The matrix A^* is the matrix of coefficients for the constraints (2.3-32) and (2.3-33). The jk^{th} component of the vector defined by (2.3-39) is

$$z_{jk} - \tilde{b}_j' \pi_{jk}^* = \underline{c}_B (B^{-1})_{1;n_1} \tilde{A}' \pi_{jk}^* + \underline{c}_B (B^{-1})_{;n_1+j}$$

$$- \tilde{b}_j' \pi_{jk}^*$$

$$= (\underline{c}_B (B^{-1})_{1;n_1} \tilde{A}' - \tilde{b}_j') \pi_{jk}^* + \underline{c}_B (B^{-1})_{;n_1+j}. \quad (2.3-40)$$

Now to find the minimum $z_{jk} - \tilde{b}_j' \pi_{jk}^*$ over all j and k , we notice that

$$\min_{j,k} (z_{jk} - \tilde{b}_j' \pi_{jk}^*) = \min_j [\min_k (z_{jk} - \tilde{b}_j' \pi_{jk}^*)] \quad (2.3-41)$$

and for any given j the minimum of $z_{jk} - \tilde{b}_j' \pi_{jk}^*$ as defined in (2.3-40) occurs at an extreme point of S_j such that

$$z_j^* = \min_{\pi_j} z_j = \min_{\pi_j} [(\underline{c}_B (\mathbb{B}^{-1})_{1;n_1} \tilde{A}' - \tilde{b}_j') \pi_j] \quad (2.3-42)$$

subject to

$$\tilde{B}' \pi_j \leq p_j \quad (2.3-43)$$

and

$$\pi_j \geq 0 \quad (2.3-44)$$

We can find the value of z_j^* as

$$z_j^* = - \max_{\pi_j} [-z_j] = - \max_{\pi_j} [(\tilde{b}_j' - \underline{c}_B (\mathbb{B}^{-1})_{1;n_1} \tilde{A}') \pi_j] \quad (2.3-45)$$

subject to (2.3-43) and (2.3-44). If the extreme point corresponding to the value z_j^* is π_{jr}^* , then the minimum in (2.3-41) becomes

$$\begin{aligned} \min_{j,k} (z_{jk} - \tilde{b}_j' \pi_{jk}^*) &= \min_j [z_j^* + \underline{c}_B (\mathbb{B}^{-1})_{;n_1+j}] \\ &= z_s^* + \underline{c}_B (\mathbb{B}^{-1})_{;n_1+s} \end{aligned} \quad (2.3-46)$$

with the corresponding extreme vector $\underline{\pi}_{sr}^*$.

The second step in the algorithm is to check for an optimal solution. If all $z_j^* + \underline{c}_B(B^{-1})_{;n_1+j}$ and all the elements of $\underline{c}_B(B^{-1})_{1;n_1}$ are non-negative then an optimal solution has been reached and we proceed with step five. If the solution is not optimal then the variable corresponding to the minimum of these values is selected as the new basic variable, say λ_{sr} , and we proceed with step three.

The third step is to find the variable which leaves the basis. In order to do this, it is necessary to know $A_{;s}^* = (\underline{d}_{sr}, \underline{e}_s)'$ where

$$\underline{d}_{sr} = \tilde{A}' \underline{\pi}_{sr}^* \quad (2.3-47)$$

and \underline{e}_s is the s-coordinate vector of dimension N. From (2.3-36) $\underline{t}^* = (\underline{c}, \underline{l}_N)'$, where \underline{c} is as given in (2.3-6) and \underline{l}_N is an N-dimensional unit vector. The basic variable which leaves the basis is found to be the variable corresponding to the minimum element of the vector

$$\underline{v} = (B^{-1} \underline{t}^*) \div (B^{-1} A_{;s}^*) , \quad (2.3-48)$$

where the minimum is taken over all elements in \underline{v} corresponding to $(B^{-1} A_{;s}^*)_i \geq 0$. If $\min \underline{v} = v_q$, then the qth basic variable is replaced by λ_{sr} .

The matrix B^{-1} and the vector \underline{c}_B must be updated in the fourth step so that

$$(\underline{c}_B)_q = \tilde{b}_s' \underline{\pi}_{sr}^* \quad (2.3-49)$$

and

$$\underline{B}_{\text{NEW}}^{-1} = E_{n_1+N}^q \underline{B}_{\text{OLD}}^{-1}, \quad (2.3-50)$$

where the matrix $E_{n_1+N}^q$ is an n_1+N by n_1+N identity matrix whose elements in the q^{th} column are

$$(E_{n_1+N}^q)_{iq} = \begin{cases} - (\underline{B}^{-1} A_s^*)_i \div (\underline{B}^{-1} A_s^*)_q & i \neq q \\ 1 \div (\underline{B}^{-1} A_s^*)_q & i = q \end{cases}, \quad (2.3-51)$$

$$i=1, 2, \dots, n_1 + N.$$

Once the transformation is complete, the optimality criterion is again checked by returning to the first step of the decomposition algorithm (2.3-39).

The fifth and final step is to transform the answer to the dual back to the primal variable \underline{x} , i.e.,

$$\underline{x}^* = \underline{c}_B \underline{B}^{-1} \quad (2.3-52)$$

and to find the optimal value of the objective function,

$$z = \underline{c}_B \underline{B}^{-1} \underline{t}^*. \quad (2.3-53)$$

To find \underline{y}_i , the linear programs

$$\min_{\underline{y}_i} (\underline{f}' \underline{y}_i) + \underline{c}' \underline{x}^* \quad (2.3-54)$$

subject to

$$B \underline{y}_i = \underline{b}_i - A \underline{x}^* \quad (2.3-55)$$

and

$$\underline{y}_i \geq 0 \quad (2.3-56)$$

for $i=1, 2, \dots, N$ must be solved.

It is obvious that a great deal of work is required to solve such a stochastic model and that for small problems it is probably better to solve the model (2.3-6), (2.3-7), and (2.3-8) and reserve the dual formulation for large problems. In any event it should be clear by now that stochastic models should not be used indiscriminately and Section 2.5 will help to clarify the particular class of problems in which they should be used.

2.4 Example of Two-Stage Linear Programming

Suppose there is a factory which can use either a machine #1 or a machine #2 or labour or any combination of the three to produce two different products. The demand for its products is a random variable with a known distribution, but the specific value of this random variable is not known when the production policy is set. Furthermore, there is a central warehouse from which the factory can draw if its production is less than the demand, or to which it can ship its surplus if its production is greater than the demand. The shipments and the factors of production have a known unit cost so that the factory must find a policy which will minimize its expected total cost subject to the demand constraints.

This problem could be represented as

$$\min_{\underline{x}, \underline{y}} E [\underline{c}' \underline{x} + \underline{f}' \underline{y}]$$

$$= \min E \left\{ (\underline{c}_1, \underline{c}_2, \underline{c}_3) \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \underline{x}_3 \end{bmatrix} + (\underline{f}_1, \underline{f}_2) \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix} \right\}$$

subject to

$$A \underline{x} + B \underline{y} = \underline{b} ,$$

or equivalently

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \underline{x}_3 \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} ,$$

and

$$\underline{x} \geq \underline{0} , \quad \underline{y} \geq \underline{0} ,$$

where b_i is the number of units demanded of the i^{th} product for $i=1,2$, and x_j , $j=1,2,3$ are the number of units of the j^{th} input factor (machine #1, labour, and machine #2, respectively) at the unit cost of c_j . The quantity a_{ij} is the number of units of the i^{th} product produced by a unit of the j^{th} input factor. The second term in the constraint is the warehouse-shipment rate where B_{il} is the number of units of the i^{th} product per shipment to the warehouse and B_{i2} is the rate of shipment in the opposite direction. The decision

variables y_1 and y_2 are, respectively, the number of shipments to and from the warehouse at a unit cost f_1 and f_2 , respectively.

If we assign some integral values to c , f , A , and B , we obtain the following sample problem:

$$z = \min_{\underline{x}, \underline{y}} E [4x_1 + 12x_2 + 9x_3 + 8y_1 + 10y_2]$$

subject to

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} -1 & 1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

and

$$x_i \geq 0, \quad i=1,2,3$$

$$y_j \geq 0, \quad j=1,2.$$

Suppose now that b can have one of two values $b_1 = (30, 45)$ or $b_2 = (36, 54)$, with probability $p_1 = 0.25$ and $p_2 = 0.75$ respectively. The factory must select the levels of x_i 's before observing b , and then compensate with the selection of y_j 's. This problem may be solved by using the method in Section 2.3. The optimal solution is $z = \$224.50$, $x_1 = 1$ hour, $x_2 = 16$ hours, and $x_3 = 0$ hours. The optimum strategy for \underline{y} is given in Table 2-1.

TABLE 2-1Second-Stage Strategy for \underline{y}

Demand	Number of Shipments	
	to warehouse	from warehouse
(30, 45)	3	0
(36, 54)	0	3

This strategy implies that if \underline{b}_1 occurs then there is a surplus which is shipped to the warehouse and the resultant cost is \$220. If the demand \underline{b}_2 occurs then there is a shortage, and the factory must bring in additional supplies from the warehouse and the resultant cost is \$226.

2.5 Inequalities for Linear Programming Under Uncertainty

In this section we shall discuss a few alternative models for the linear programming problem under uncertainty. These alternative models will then be used to show the influence of uncertainty in linear programs. Madansky (1960) was the first to use this type of analysis and a good summary of his work can be found in Wagner (1969).

The formulation of the stochastic model in Section 2.3 is referred to as the "here-and-now" approach since a decision

must be made before all of the information about the system is known. For convenience, let us denote the objective function by

$$\min_{\underline{x}} E [C(\underline{b}, \underline{x})] \equiv E[\text{here and now}] \quad (2.5-1)$$

where

$$C(\underline{b}, \underline{x}) = \min_{\underline{y}} (\underline{c}' \underline{x} + \underline{f}' \underline{y}) . \quad (2.5-2)$$

The following analysis depends on the assumption that for any given \underline{x} and \underline{b} there always exists a feasible solution \underline{y} to the here-and-now model. The distribution of the random elements, \underline{b} , need not be discrete for the following analysis to apply.

If a decision on the variables \underline{x} , and \underline{y} is postponed until the random elements are observed then the model is called a "wait-and-see" model. The problem then is to find the

$$\min_{\underline{x}, \underline{y}} (\underline{c}' \underline{x} + \underline{f}' \underline{y}) \quad (2.5-3)$$

subject to (2.3-2) and (2.3-4) where \underline{b} is known. The expected value of the optimum solution for this model is

$$E[\min_{\underline{x}, \underline{y}} (\underline{c}' \underline{x} + \underline{f}' \underline{y})] \equiv E[\text{wait and see}] . \quad (2.5-4)$$

If the random elements have a discrete distribution, it would be possible to solve every possible combination for the

optimum value, weight these values with the corresponding probabilities, and then calculate the expected value. This would involve solving deterministic linear programs of the form

$$\min_{\underline{x}, \underline{y}} (\underline{c}' \underline{x} + \underline{f}' \underline{y}) \quad (2.5-5)$$

subject to

$$A \underline{x} + B \underline{y} = b_i \quad (2.5-6)$$

and

$$\underline{x} \geq 0, \quad \underline{y} \geq 0, \quad (2.5-7)$$

where $i=1, 2, \dots, N$.

Consider, for example, the problem in Section 2.4. If all possible problems were formed and solved, then

$$\begin{aligned} E[\text{wait and see}] &= \frac{1}{4} (\$180) + \frac{3}{4} (\$216) \\ &= \$207. \end{aligned}$$

Note that $E[\text{here and now}] = \224.50 and hence

$$E[\text{wait and see}] < E[\text{here and now}] . \quad (2.5-8)$$

Madansky (1960) has shown that (2.5-8) is generally true. Since the wait-and-see model is not subject to uncertainties, it is a system of complete information, and thus it is reasonable that the expected minimum total cost is less.

If the average values of the random elements are used in place of the random elements then the model is called the "average-value" model and may be stated as follows:

$$\min_{\underline{x}, \underline{y}} (\underline{c}' \underline{x} + \underline{f}' \underline{y}) \equiv E[\text{average value}] \quad (2.5-9)$$

subject to

$$A \underline{x} + B \underline{y} = E[\underline{b}] \quad (2.5-10)$$

and

$$\underline{x} \geq \underline{0}, \quad \underline{y} \geq \underline{0}. \quad (2.5-11)$$

If we consider the example of the previous section, then the optimum value for the average-value problem is $E[\text{average value}] = \207 , which is equal to the solution to the wait-and-see model. Madansky (1960) found that in general

$$E[\text{average value}] \leq E[\text{wait and see}] \quad (2.5-12)$$

and in particular states that:

"if the probability measure on the set of \underline{b} 's is countably additive or the set of \underline{b} 's is with probability one finite in number, equality holds between

$$E[\text{wait and see}] = E[\text{average value}] \quad (2.5-13)$$

if and only if $\min_x C(\underline{b}, \underline{x})$ is a linear function of \underline{b} ."

If the optimal feasible value of \underline{x} in the average-value model is \underline{x}^* , then the model obtained by using \underline{x}^* as the first-stage strategy and \underline{y} as the second-stage strategy is called the "first-stage, average-value approximation". Mathematically this means that

$$\min_{\underline{x}, \underline{y}} (\underline{c}' \underline{x} + \underline{f}' \underline{y}) = \underline{c}' \underline{x}^* + \min_{\underline{y}} (\underline{f}' \underline{y}) \quad (2.5-14)$$

subject to

$$B \underline{y} = \underline{b} - A \underline{x}^* \quad (2.5-15)$$

and

$$\underline{y} \geq 0 \quad . \quad (2.5-16)$$

Hence,

$$E[\underline{c}' \underline{x}^* + \min_{\underline{y}} (\underline{f}' \underline{y})] \equiv E[\text{first-stage, average-value approx.}] . \quad (2.5-17)$$

Again the expected value of this approximation can be found simply for a discrete distribution function by solving all possible problems as was done in the wait-and-see model. In the example, the first problem would be to find

$$\min_{\underline{y}} (8, 10) \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

subject to

$$\begin{bmatrix} -1 & 1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -4.5 \\ -6.75 \end{bmatrix}$$

and

$$y_1 \geq 0, \quad y_2 \geq 0.$$

Clearly, there is no feasible solution to this problem since it violates the assumption that there always exists a \underline{y} which is feasible for each pair $(\underline{b}, \underline{x})$.

It may be shown that, the expected value of the first-stage, average-value approximation, if it exists, is an upper bound for the expected value of the here-and-now method, i.e.,

$$E[\text{here and now}] \leq E[\text{first-stage, average-value approx.}] . \quad (2.5-18)$$

There is an interesting economic interpretation to the difference

$$\begin{aligned} E[\text{here and now}] - E[\text{wait and see}] \\ = \$224.50 - \$207 = \$17.50 \end{aligned} \quad (2.5-19)$$

in that it reflects the cost of uncertainty. This is the extra cost due to choosing \underline{x} prior to observing the random demand \underline{b} . It is also the amount the firm would be willing to pay to be able to predict the demand perfectly. Suppose that the difference

$$E[\text{first-stage, average-value approx.}] - E[\text{average value}] \quad (2.5-20)$$

exists where both of expectations are simple calculations. If this difference is "small" then it may be more economical to use \underline{x}^* as the first-stage optimal decision variable for the here-and-now model. This plan will be more economical if the extensive form of the here-and-now model is very large. Thus the difference (2.5-20) indicates the closeness of the optimal values for the average-value and the here-and-now models.

CHAPTER 3

CHANCE-CONSTRAINED LINEAR PROGRAMMING

3.1 Formulation of the Model

In this chapter we shall consider an alternative model for the classical linear programming problem in which the coefficients are subject to stochastic variations. It was developed originally by Charnes, Cooper, and Symonds (1958) in considering the scheduling of heating oil production, the demand for which is subject to marketing and storage constraints. This model, sometimes called the E-model, may be stated in simplified form as

$$\min_{\underline{x}} \quad E[\underline{c}' \underline{x}] \quad (3.1-1)$$

subject to

$$\Pr(\underline{A} \underline{x} \geq \underline{b}) \geq \underline{\alpha}, \quad (3.1-2)$$

or

$$\Pr\left(\sum_{j=1}^n a_{ij} x_j \geq b_i\right) \geq \alpha_i, \quad i=1, 2, \dots, m,$$

where

$$0 \leq \alpha_i \leq 1, \quad i=1, 2, \dots, m,$$

and

$$\underline{x} \geq \underline{0}. \quad (3.1-3)$$

The deterministic constraints of the classical model are replaced by the chance constraints (3.1-2). In the solution of this model, which will be represented in Section 3.2, the chance constraints are replaced by equivalent deterministic constraints using the distribution function of \underline{b} . Charnes and Cooper (1959) give a more general and theoretical treatment of the above type of problem.

The problem of scheduling an oil tanker fleet is considered by Charnes and Cooper (1962). An optimal ship-chartering policy is found subject to an uncertain shipping demand. The cost of chartering ships is assumed to be a random variate which depends on a normally distributed shipping demand. As in the heating oil problem, they are interested in a scheduling problem, consequently they introduce a linear decision rule into the model (see Section 3.3).

Charnes and Cooper (1963 a) consider, as well as the E-model, some alternative forms of the objective function (3.1-1). If the objective function is replaced by

$$\min_{\underline{x}} E[(\underline{c}' \underline{x} - \underline{c}'_0 \underline{x}_0)^2], \quad (3.1-4)$$

where \underline{c}'_0 and \underline{x}_0 are given vectors, the model is called the V-model. The model formed by replacing the objective function by

$$\max_{\underline{x}} \Pr(\underline{c}' \underline{x} \geq \underline{c}' \underline{x}_0) \quad (3.1-5)$$

is called the P-model.

Miller and Wagner (1965), and Symonds (1967, 1968) consider replacing the m constraints (3.1-2) by a single joint constraint

$$\Pr(\underline{A} \underline{x} \geq \underline{b}) \geq \alpha . \quad (3.1-6)$$

This model is simpler since only one value of α has to be specified, but the solution is more complicated. Let the cumulative distribution function of b_i be

$$F_i(\eta_i) = \Pr(b_i \leq \eta_i), \quad \eta_i \in \mathbb{R}, \quad (3.1-7)$$

where F_i is continuous and the b_i , $i=1, 2, \dots, m$, are assumed to be independent. The joint distribution of \underline{b} can then be written as

$$\Pr(\underline{b} \leq \underline{\eta}) = F_{\underline{b}}(\underline{\eta}) = \prod_{i=1}^m F_i(\eta_i) . \quad (3.1-8)$$

Then the equivalent deterministic constraint for (3.1-6) is

$$\underline{A} \underline{x} \geq \underline{\eta} , \quad (3.1-9)$$

where

$$\prod_{i=1}^n F_i(\eta_i) \geq \alpha , \quad (3.1-10)$$

with the sign of η_i unrestricted. The function $F_{\underline{b}}(\eta)$ is nonlinear and, according to Wagner (1969), is rarely convex. Thus nonlinear programming can rarely be used as the problem in its present form. However, if we take the natural logarithm of both sides of (3.1-10), i.e.,

$$\sum_{i=1}^m \ln F_i(\eta_i) \geq \ln \alpha , \quad (3.1-11)$$

then the left-hand side is a convex function when b_i is a normal, gamma, or uniform variate. Consequently, the joint chance-constrained program can usually be solved using nonlinear techniques.

3.2 Solution of the Chance-Constrained Model

The chance-constrained linear programming model may be used in place of the classical linear programming model when it is desirable but not absolutely essential that the constraints hold. The α_i in (3.1-2) are specified constants in the interval $[0,1]$ and are usually chosen close to one. The quantity $1-\alpha_i$ represents the allowable risk that x_i will not satisfy (3.1-2) once the values of the random variables are observed, i.e.,

$$\sum_{j=1}^n a_{ij} x_j < b_i . \quad (3.2-1)$$

Thus it is necessary in the chance-constrained program to define the vector \underline{x} as feasible if and only if \underline{x} satisfies

(3.1-2) and (3.1-3). The objective is to select the "best" nonnegative solution that will "probably" satisfy the constraints $A \underline{x} \geq \underline{b}$ after the values of the random variables A , \underline{b} and \underline{c} are observed.

Let us consider a particular subset of the class of chance-constrained programs where a_{ij} and α_i are known constants. The variables c_j and b_j are random and statistically independent for all i and j . Again the distribution function, $F_i(b) = \Pr(b_i \leq b)$, for each b_i is known and is assumed to be normal with mean μ_i and standard deviation σ_i . To solve this specific chance-constrained linear program, the probability relations must be replaced by equivalent deterministic constraints.

If y_i is the α_i -fractile of F_i , i.e.,

$$\alpha_i = F_i(y_i) = \int_{-\infty}^{y_i} \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{(b_i - \mu_i)^2}{2\sigma_i^2}} db_i , \quad (3.2-2)$$

then

$$\Pr(b_i \leq y_i) = \Pr(y_i \geq b_i) = F_i(y_i) = \alpha_i . \quad (3.2-3)$$

Since F_i is a continuous monotonically increasing function in y_i , the relation

$$\Pr(\sum_{j=1}^n a_{ij} x_j \geq b_i) \geq \alpha_i \quad (3.2-4)$$

is equivalent to

$$\sum_{j=1}^n a_{ij} x_j \geq F_i^{-1}(\alpha_i) = y_i . \quad (3.2-5)$$

However, the right-hand side of this last inequality would be more useful if it were expressed in terms of a standar-dized normal variate, i.e.,

$$F_i^{-1}(\alpha_i) = y_i = \sigma_i \Phi^{-1}(\alpha_i) + \mu_i , \quad (3.2-6)$$

where

$$\alpha_i = \Phi(z_i) = \int_{-\infty}^{z_i} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx . \quad (3.2-7)$$

Therefore the i^{th} chance-constraint in (3.1-2) can be replaced by

$$\sum_{j=1}^n a_{ij} x_j \geq \sigma_i \Phi^{-1}(\alpha_i) + \mu_i , \quad i=1,2,\dots,m , \quad (3.2-8)$$

and the chance-constrained program defined previously in Section 3.1 is equivalent to the classical program

$$\min_{\underline{x}} \sum_{j=1}^n E[c_j] x_j \quad (3.2-9)$$

subject to (3.2-8) and $x_j \geq 0$, $j=1,2,\dots,n$.

Now let us consider a chance-constrained model for the example introduced in Section 2.4, viz.,

$$\min_{\underline{x}_i} [4x_1 + 12x_2 + 9x_3]$$

subject to

$$\Pr(x_1 + 2x_2 + x_3 \geq b_1) \geq 0.95$$

$$\Pr(3x_1 + 3x_2 + x_3 \geq b_2) \geq 0.95 ,$$

and

$$x_1, x_2, x_3 \geq 0 .$$

If b_1 and b_2 are normal variates with means 34.5 and 51.75, respectively, and standard deviations 1.0 and 1.5, respectively, then the equivalent classical model is

$$\min_{\underline{x}} \left\{ (4, 12, 9) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right\}$$

subject to

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \geq \begin{bmatrix} 1.0(1.645) + 34.5 \\ 1.5(1.645) + 51.75 \end{bmatrix} = \begin{bmatrix} 35.145 \\ 54.2175 \end{bmatrix} ,$$

and

$$\underline{x} \geq 0 ,$$

and $\Phi^{-1}(0.95) = 1.645$. The solution of this linear programming problem is $x_1 = 1$ hr. for machine #1, $x_2 = 17.0725$

hrs. of labour, and $x_3 = 0$ hrs. for machine #2, and the minimum value of the objective function is approximately \$208.87.

3.3 Linear Decision Rules

As mentioned in Section 3.1, Charnes and Cooper (1958, 1959, 1962, 1963a) consider the problem of policy planning over several time periods. If x_j represents the production rate during the j^{th} period, then the problem is to select the x_j so that the total cost is minimized over the entire planning period. The scheduling problem can be incorporated into the simple chance-constrained program (3.1-1), (3.1-2), and (3.1-3) if the notion of a linear decision rule is introduced. A linear decision rule for x_j is a set of constants d_{jk} and a decision variable y_j such that

$$x_j = \sum_{k=1}^m d_{jk} b_k + y_j, \quad j=1, 2, \dots, n. \quad (3.3-1)$$

If b_k is not known until after a value is assigned to x_j , then $d_{jk} = 0$. Hence, $\underline{x} = D \underline{b} + \underline{y}$ is a linear decision rule for \underline{x} , if the matrix D is a lower triangular matrix of the form

$$D = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ d_{21} & 0 & 0 & \dots & 0 & 0 \\ d_{31} & d_{32} & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ d_{n1} & d_{n2} & d_{n3} & \dots & d_{nm-1} & 0 \end{bmatrix} \quad (3.3-2)$$

and b_1, b_2, \dots, b_{j-1} are known before a value is assigned to x_j . The quantity d_{jk} can be considered as the throughput rate since it is the ratio of input to output if \underline{x} and \underline{b} are, respectively, the input and output vectors of the production process.

If the linear decision rule (3.3-1) is substituted for \underline{x} in the objective function (3.1-1), then the resulting objective function is of the form

$$\begin{aligned} \min_{\underline{x}} E[\underline{c}' \underline{x}] &= \min_{\underline{y}} E[\underline{c}' (D \underline{b} + \underline{y})] \\ &= \min_{\underline{y}} \{E[\underline{c}' D \underline{b}] + E[\underline{c}' \underline{y}]\} \\ &= E[\underline{c}'] D E[\underline{b}] + \min_{\underline{y}} (E[\underline{c}'] \underline{y}) . \quad (3.3-3) \end{aligned}$$

Since the first term in (3.3-3) is a constant, only the second term is relevant to the new problem. When the linear decision rule is substituted into (3.1-2), the resulting probability relation is

$$\begin{aligned} \Pr(A \underline{x} \geq \underline{b}) &= \Pr[A(D \underline{b} + \underline{y}) \geq \underline{b}] \\ &= \Pr(A D \underline{b} + A \underline{y} \geq \underline{b}) \\ &= \Pr(A \underline{y} \geq \underline{b} - A D \underline{b}) \geq \alpha . \quad (3.3-4) \end{aligned}$$

Recall that \underline{b} is a normal variate with mean vector μ and variance-covariance matrix Σ , $N(\mu, \Sigma)$. If

$\underline{\eta} = \underline{b} - A D \underline{b}$ then $\underline{\eta}$ is $N(E[\underline{b}] - A D E[\underline{b}], E[\underline{\eta} \underline{\eta}'])$

or $N(\mu_{\underline{\eta}}, \Sigma_{\underline{\eta}})$, where

$$E[\eta_i] = \mu_{\eta_i} = \mu_i - \sum_{j=1}^n \sum_{k=1}^m a_{ij} d_{jk} \mu_k, \quad (3.3-5)$$

and

$$\text{Var}(\eta_i) = \sigma_{\eta_i}^2 = \sum_{\substack{k=1 \\ k \neq i}}^m \left\{ \sum_{j=1}^n a_{ij} d_{jk} \right\}^2 \sigma_k^2 + \left\{ 1 - \sum_{j=1}^n a_{ij} d_{ji} \right\}^2 \sigma_i^2 \quad (3.3-6)$$

for $i=1, 2, \dots, m$. Since the distribution of η_i is known, and the problem ((3.3-3) and (3.3-4)) satisfies the assumptions of Section 3.2 the deterministic equivalent to (3.3-4) can be found. Accordingly,

$$\Pr \left(\sum_{j=1}^n a_{ij} x_j \geq b_i \right) \geq \alpha_i, \quad (3.3-7)$$

if and only if

$$\sum_{j=1}^n a_{ij} y_j \geq \mu_{\eta_i} + \Phi^{-1}(\alpha_i) \sigma_{\eta_i}, \quad i=1, 2, \dots, m. \quad (3.3-8)$$

In summary, the use of a linear decision rule (3.3-1) results in the classical program defined by (3.3-3), (3.3-8) and $y_j \geq 0, j=1, 2, \dots, n$.

3.4 Discussion

The major disadvantage of the chance-constrained program is that it requires more simplifying assumptions

than the two-stage linear program under uncertainty.

Chance-constrained programming only indirectly evaluates the economic consequences of violating a constraint, and it makes no distinction between a small and a large shortage of production. Also the values α_i of the chance-constrained model must be specified before a solution can be found.

However, the chance-constrained formulation does have some advantages. The equivalent classical program for the chance-constrained model has the same number of constraints as the original program. Whereas the greatest handicap of the two-stage linear program under uncertainty (for discrete b) is the resultant size of its equivalent classical program. Further, the solution of the chance-constrained model is much simpler than the complete strategy solution for a linear program under uncertainty. Finally, the chance-constrained model requires less information about the distribution of b_i - only the α_i -fractile of b_i need be known. Therefore these advantages and disadvantages must be weighed carefully when deciding on a particular model for a given problem.

CHAPTER 4

SYSTEMS OF STOCHASTIC LINEAR EQUATIONS

4.1 Introduction

A stochastic linear programming model has the form

$$z = \max_{\underline{x}} \underline{c}' \underline{x} \quad (4.1-1)$$

subject to

$$A \underline{x} \leq \underline{b} \quad (4.1-2)$$

and

$$\underline{x} \geq 0 , \quad (4.1-3)$$

where A is an m by n matrix of random variables, \underline{b} is an m -dimensional vector of random variables, and \underline{c} is an n -dimensional vector of random variables. The joint probability density function of A , \underline{b} , and \underline{c} ,

$$\Pr(A, \underline{b}, \underline{c} | \underline{\theta}) , \quad (4.1-4)$$

is assumed to be known and may depend upon a given parameter $\underline{\theta}$. The purpose of stochastic programming is to investigate the distribution of z by solving for its density function or its moments. Also if there exist several possible probability density functions,

$$\Pr(A, \underline{b}, \underline{c} | \underline{\theta}) , \quad \underline{\theta} \in \Omega , \quad (4.1-5)$$

the stochastic programming may be used to select the "best" probability density function, defined to be that density function which maximizes with respect to $\underline{\theta}$ some given preference functional, e.g.,

$$\max_{\underline{\theta} \in \Omega} E [z | \underline{\theta}] . \quad (4.1-6)$$

Instead of the expectation, the mode, median, variance, skewness, or some other statistic of the distribution of z may be used in (4.1-6). Ideally, the solution, i.e., the density of z , should be derived analytically, but as will be seen this is not practical for the majority of problems. Consequently, we must settle for some approximate solution.

Suppose that the slack variables have already been added to the above model so that the constraints (4.1-2) may be rewritten as

$$A \underline{x} = \underline{b} . \quad (4.1-7)$$

If the simplex algorithm were applied to the deterministic model defined by (4.1-1), (4.1-7), and (4.1-3), then the simplex algorithm would select a submatrix of A , say A_1 , consisting of m columns so that the objective function $\underline{c}' \underline{x} = \underline{c}_1' \underline{x}_1 + \underline{c}_2' \underline{x}_2$ is maximized and where \underline{x}_1 and \underline{x}_2 are such that

$$A_1 \underline{x}_1 = \underline{b} \quad (4.1-8)$$

and

$$\underline{x}_2 = \underline{0} . \quad (4.1-9)$$

The subvectors \underline{x}_1 and \underline{c}_1 correspond to the particular submatrix A_1 . Thus in the simplex algorithm we must ultimately solve the set of linear equations (4.1-8). Correspondingly, in the stochastic linear program we must solve a system of stochastic linear equations for the distribution of \underline{x} , where

$$A \underline{x} = \underline{b} . \quad (4.1-10)$$

In (4.1-10) A is assumed to be an m by m matrix and \underline{b} an m -dimensional vector both composed of independent random variables. It is the purpose of this chapter to investigate such systems of stochastic linear equations while the next chapter will discuss stochastic linear programming.

One of the first attempts to find an approximate solution to a system of stochastic linear equations, in which the elements of A and \underline{b} were assumed to be independent normal variates, was by Babbar (1955). His approach is based on solving for \underline{x} using Cramer's rule and then expanding the required determinants in terms of their cofactors. He then ignores for no apparent reason all resulting second and higher order terms of random variables. Thus he is able to express the

components of \underline{x} as a ratio of two linear combinations of normal variates. (He claims that the ratio of two normal variates is approximately normally distributed if certain criteria are satisfied.) He then is able to approximate the distribution of this ratio by transforming this quantity into a variate which is approximately normal. It will be shown in the next section that this ratio of two normal variates can be solved analytically, hence no approximation is necessary. Babbar (1955) also extends this method to stochastic linear programming by assuming that only one basis may be optimal, i.e., the chance of some other basis being optimal is assumed to be zero. His approach to the problem has been criticized by Wagner (1958) as being restrictive and vacuous because he ignores the non-negativity requirements.

4.2 Transformation of Random Variables

There does not exist any literature pertaining directly to the analytical solution of the stochastic linear equations of the form (4.1-10). Most of the literature relates to systems of linear equations which are subject to stochastic errors and to the determination of error bounds for \underline{x} . This type of error analysis is of no help when \underline{x} is a random variable for which we

require a probability density function or moments.

However, we can find an analytic solution to a system of stochastic linear equations by applying the theory of random variable transformations (see Hogg and Craig (1965)).

The system of stochastic linear equations

$$A \underline{x} = \underline{b}, \quad (4.2-1)$$

may be viewed as a transformation of random variables.

The equations (4.2-1) represent a transformation of R^{m^2+m} to R^m and we have assumed that the joint probability density function of A and \underline{b} ,

$$\Pr(A, \underline{b}) = \Pr(a_{11}, a_{12}, \dots, a_{mm}, b_1, \dots, b_m) = \phi(\underline{z}), \quad (4.2-2)$$

is known. If we introduce m^2 additional variables $\underline{y} = (y_1, \dots, y_{m^2})$ as a function of A and \underline{b} , e.g.,

$$\begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix} = \begin{bmatrix} A^{-1} & O \\ \vdots & \vdots \\ a_{11} & 0 & \dots & 0 \\ \vdots & 0 & a_{12} & \dots & 0 \\ O & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ddots & a_{mm} & 1 \end{bmatrix} \begin{bmatrix} \underline{b} \\ \vdots \\ 1 \end{bmatrix} \quad (4.2-3)$$

or equivalently

$$\tilde{\underline{x}} = \tilde{A}^{-1} \tilde{\underline{b}}, \quad (4.2-4)$$

and select these additional variables so that the mapping is continuous, one to one, and onto then the inverse mapping exists. Suppose it is possible to express the inverse mapping, $\tilde{\underline{b}} = \tilde{A} \tilde{\underline{x}}$, as

$$z_i = \tilde{b}_i = w_i(\tilde{\underline{x}}), \quad i=1,2,\dots,m, \quad (4.2-5)$$

and

$$z_k = \tilde{a}_{ij} = w_k(\tilde{\underline{x}}), \quad k = mi + j \quad \text{and} \quad (4.2-6) \\ i,j=1,2,\dots,m.$$

If the first-order partial derivatives of the inverse functions are continuous, then we may express the Jacobian, J , of this transformation as

$$J = \det(U), \quad u_{ij} = \frac{\partial z_i}{\partial x_j}, \quad i,j=1,2,\dots,m^2+m. \quad (4.2-7)$$

The joint probability density of \underline{x} may be expressed as

$$g(\tilde{\underline{x}}) = |J| \phi(w_1(\tilde{\underline{x}}), \dots, w_{m^2+m}(\tilde{\underline{x}})) \quad (4.2-8)$$

provided that the Jacobian is non-zero on \mathbb{R}^{m^2+m} . Finally, the joint probability density function of \underline{x} can be found from (4.2-8) as

$$g_1(\underline{x}) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} g(\tilde{\underline{x}}) dy_1 \dots dy_{m^2+m}. \quad (4.2-9)$$

This method while very indirect provides the only general method of achieving an analytic solution to the system of stochastic linear equations. In special cases the transformation approach may be shortened by applying theorems on linear combinations of random variables, and these particular cases are noted in Section 4.5. It should be obvious, that the transformation approach is impractical because of its complexity, especially for problems involving more than $m^2 + m = 2$ dimensions.

We will illustrate the above procedure with a simplification of Babbar's example, the ratio of normal variates. Suppose that $A = a$ is a normal variate with mean zero and standard deviation σ_1 and $b = b$ is also a normal variate with mean zero and standard deviation σ_2 . We also assume that a and b are independent. Then the problem is to find the density of

$$x = b/a. \quad (4.2-10)$$

We will include the additional variable $y = a$ to (4.2-10) to obtain a continuous, one-to-one, and onto mapping. The inverse of this transformation is

$$\begin{aligned} a &= y \\ b &= x y \end{aligned} \quad (4.2-11)$$

with the Jacobian

$$J = \begin{vmatrix} 0 & 1 \\ y & x \end{vmatrix} = -y . \quad (4.2-12)$$

The joint probability density of a and b is

$$f(a, b) = \frac{1}{2\sigma_1\sigma_2\pi} \exp \left\{ -\frac{a^2\sigma_2^2 + b^2\sigma_1^2}{2\sigma_1^2\sigma_2^2} \right\} . \quad (4.2-13)$$

If ϕ in (4.2-8) is replaced by (4.2-13), then the joint probability density function of x and y is

$$g(x, y) = |J| f(y, xy) . \quad (4.2-14)$$

The marginal density of g is the probability density function of x, i.e.,

$$g_1(x) = \int_{-\infty}^{+\infty} |J| f(y, xy) dy \quad (4.2-15)$$

or

$$g_1(x) = \frac{1}{2\sigma_1\sigma_2\pi} \int_{-\infty}^{+\infty} |-y| \exp \left\{ -\frac{y^2\sigma_2^2 + x^2y^2\sigma_1^2}{2\sigma_1^2\sigma_2^2} \right\} dy . \quad (4.2-16)$$

Removing the absolute value operator from (4.2-16), we find that

$$g_1(x) = \frac{1}{\sigma_1\sigma_2\pi} \int_0^{\infty} y \exp \left\{ -\frac{y^2\sigma_2^2 + x^2y^2\sigma_1^2}{2\sigma_1^2\sigma_2^2} \right\} dy . \quad (4.2-17)$$

Replacing the variable y by

$$y = \frac{\sigma_1 \sigma_2 \sqrt{2u}}{\sqrt{x^2 \sigma_1^2 + \sigma_2^2}}, \quad (4.2-18)$$

the density g_1 becomes

$$g_1(x) = \frac{\sigma_1 \sigma_2}{\pi(x^2 \sigma_1^2 + \sigma_2^2)} \int_0^\infty e^{-u} du. \quad (4.2-19)$$

Thus the probability density function of $x = b/a$ is the Cauchy density,

$$g_1(x) = \frac{1}{\pi} \cdot \frac{\sigma_2/\sigma_1}{x^2 + (\sigma_2/\sigma_1)^2}, \quad -\infty < x < +\infty. \quad (4.2-20)$$

The following result is given by Kendall and Stuart (1958, Vol. 1 pp. 270-271) for the more general problem when a is $N(\mu_1, \sigma_1)$ and b is $N(\mu_2, \sigma_2)$:

$$g_1(x) = \frac{1}{\sqrt{2\pi}} \cdot \frac{\mu_1 \sigma_2^2 + \mu_2 \sigma_1^2 x}{(\sigma_2^2 + \sigma_1^2 x^2)^{3/2}} \exp\left\{-\frac{(\mu_2 - \mu_1 x)^2}{2(\sigma_2^2 + \sigma_1^2 x^2)}\right\}. \quad (4.2-21)$$

From this example we can see that even for the simplest problem the transformation of variables technique is very complicated. For this reason we will have to be satisfied with an approximate distribution for \underline{x} . This may be done by finding some algebraic expression which tends to the probability density function of \underline{x} under given conditions or by finding a relative frequency distribution for \underline{x} via simulation. Even using such approximations we

will have to be satisfied with estimates of the expected value, variance, and covariance of \underline{x} .

4.3 Tintner's Approximation

The most widely used approximation for \underline{x} is due to Tintner (1955). He considers a system of stochastic linear equations

$$A^* \underline{x} = \underline{b} \quad (4.3-1)$$

in which all the elements of A^* and \underline{b} are independent normal variates with known means and variances. The approximation is based on a Taylor's series expansion of \underline{x} about the means retaining only the linear terms (see Cramer (1946) p. 218). Applying the differential operator \mathcal{D} to (4.3-1), we obtain

$$A^* \mathcal{D}(\underline{x}) + \mathcal{D}(A^*) \underline{x} = \mathcal{D}(\underline{b}) . \quad (4.3-2)$$

Rewriting (4.3-2), we find that

$$\mathcal{D}(\underline{x}) = -(A^*)^{-1} \mathcal{D}(A^*) \underline{x} + (A^*)^{-1} \mathcal{D}(\underline{b}) . \quad (4.3-3)$$

From (4.3-3) and the Taylor's series expansion for \underline{x} , Tintner obtains his estimator $\hat{\underline{x}}_t$ for \underline{x} , where

$$\begin{aligned} \hat{\underline{x}}_t &= M \underline{b}_0 - M A M \underline{b}_0 + M(\underline{b} - \underline{b}_0) \\ &= M \underline{b} - M A M \underline{b}_0 , \end{aligned} \quad (4.3-4)$$

and $A = A^* - M_O$, $E[A^*] = M_O = M^{-1}$, and $E[\underline{b}] = \underline{b}_O$.

To find the expected value of \hat{x}_t , the expected value operator is applied to (4.3-4), i.e.,

$$E[\hat{x}_t] = M E[\underline{b}] - M E[A] M \underline{b}_O = M \underline{b}_O . \quad (4.3-5)$$

Since \hat{x}_t is a linear combination of normal variates A and \underline{b} , \hat{x}_t is a normal variate whose variance is a linear combination of the variances of A and \underline{b} .

The only necessary condition for the estimator \hat{x}_t is that x must have continuous first and second order derivatives in the neighbourhood of the point

$E [a_{11}, a_{12}, \dots, a_{mm}, b_1, \dots, b_m]$. There is no indication of how this condition relates to the original model. Tintner ignores the existence of the error $x - \hat{x}_t$ and the behaviour of this error. In fact, he implicitly assumes that \hat{x}_t is always a reasonable estimator for x . The two types of problems to which he has applied his approximation are either when A is random (Tintner (1955)) or \underline{b} is random (Tintner and Raghavan (1965)), but not when both are random.

Tintner (1955) considers the problem in which $A=a$ is $N(\mu_1, \sigma_1)$ and $\underline{b} = b$ is $N(\mu_2, \sigma_2)$, and using (4.3-4), he finds that

$$E[\hat{x}_t] = M \underline{b}_O = \mu_2/\mu_1 , \quad (4.3-6)$$

and

$$\begin{aligned}\text{Var}(\hat{x}_t) &= M^2 \text{Var}(b) - M^4 b_o^2 \text{Var}(a) \\ &= \frac{\mu_1^2 \sigma_2^2 + \sigma_1^2 \mu_2^2}{\mu_1^4} .\end{aligned}\quad (4.3-7)$$

Since \hat{x}_t is linear combination of a and b , \hat{x}_t is normally distributed and has the probability density function

$$g(\hat{x}_t) = \frac{\mu_1^2}{\sqrt{2\pi} (\mu_2^2 \sigma_1^2 + \mu_1^2 \sigma_2^2)^{\frac{1}{2}}} \exp \left\{ -\frac{\mu_1^4 (\hat{x}_t - (\mu_2/\mu_1))^2}{2(\mu_2^2 \sigma_1^2 + \mu_1^2 \sigma_2^2)} \right\} . \quad (4.3-8)$$

In the previous section we found the probability density function of x to be (4.2-21). The functions (4.2-21) and (4.3-6) do not agree to any great extent which again points out the weakness of Tintner's method in that there are no guide lines for using \hat{x}_t as an estimator for x .

4.4 Theorems on Matrix Algebra

In this section we will define several terms and state a number of theorems from which we can derive the results in Section 4.5. Again we will consider a system of linear equations

$$A \underline{x} = \underline{b} \quad (4.4-1)$$

where A is an m by m matrix, and \underline{x} and \underline{b} are m -dimensional vectors.

The sequence of matrices $A^{(1)}, A^{(2)}, \dots$, is said to have a limit A if each sequence of corresponding components, $a_{ij}^{(1)}, a_{ij}^{(2)}, \dots$, has a finite limit, and the components of A , a_{ij} , are the limiting values, i.e.,

$$\lim_{n \rightarrow \infty} A^{(n)} = A, \quad \text{if} \quad \lim_{n \rightarrow \infty} a_{ij}^{(n)} = a_{ij} \quad \text{for all } i \text{ and } j. \quad (4.4-2)$$

If \mathcal{D} is now used as the diagonalization operator and $\mathcal{D}(\underline{\lambda})$ is a diagonal matrix whose elements are the characteristic values of A , then we can prove the following theorem.

Theorem 2

If A can be represented in the form

$$A = V \mathcal{D}(\underline{\lambda}) V^{-1}, \quad (4.4-3)$$

where the i^{th} column of V is the i^{th} characteristic vector of A , then $A^n \rightarrow 0$ as $n \rightarrow \infty$ if and only if $\mathcal{D}^n(\underline{\lambda}) \rightarrow 0$ as $n \rightarrow \infty$ (or equivalently all the characteristic roots of A are less than unity in absolute value).

It can also be shown that if the matrix A is non-negative and has all its elements and column sums less than one, then $A^n \rightarrow 0$ as $n \rightarrow \infty$.

Theorem 3

If the matrix $A = [a_{ij}]$ is such that $0 \leq a_{ij} < 1$ for $i, j = 1, 2, \dots, m$ and

$$\sum_{i=1}^m a_{ij} < 1 \quad , \quad j = 1, 2, \dots, m , \quad (4.4-4)$$

then

$$\lim_{n \rightarrow \infty} A^n = O . \quad (4.4-5)$$

Next let us define the vector norm, $\|\underline{x}\|$, of \underline{x} as a measure having the properties:

$$(i) \quad \|\underline{x}\| \geq 0 \quad (4.4-6)$$

$$(ii) \quad \|\underline{x}\| = 0 \quad \text{then} \quad \underline{x} = 0 \quad (4.4-7)$$

$$(iii) \quad \|c \underline{x}\| = |c| \|\underline{x}\| \quad (4.4-8)$$

$$(iv) \quad \|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\| . \quad (4.4-9)$$

A matrix norm must also have the properties (i) through (iv) together with the additional property:

$$(v) \quad \|AB\| \leq \|A\| \|B\| . \quad (4.4-10)$$

The matrix norm provides an alternative criterion for $A^n \rightarrow 0$ as $n \rightarrow \infty$.

Theorem 4

If $\|A\| \leq 1$ then $A^n \rightarrow 0$ as $n \rightarrow \infty$.

Proof: We know from (v) that

$$\|A^n\| \leq \|A^{n-1}\| \|A\| , \quad (4.4-11)$$

and by repeated use of this property we conclude that

$$0 \leq \|A^n\| \leq \|A\|^n . \quad (4.4-12)$$

Since $0 \leq \|A\| < 1$, $\|A\|^n \rightarrow 0$ as $n \rightarrow \infty$, and (4.4-12) holds for all n , $\|A^n\| \rightarrow 0$. Therefore, property (ii) implies that $A^n \rightarrow 0$ as $n \rightarrow \infty$.

Q.E.D.

With the above criterion of $A^n \rightarrow 0$ as $n \rightarrow \infty$ we can now prove a theorem which will permit us to express the inverse of a matrix as a series of matrices.

Theorem 5

If $|I-A| \neq 0$ and $A^n \rightarrow 0$ as $n \rightarrow \infty$, then

$$I + A + A^2 + \dots = (I - A)^{-1} . \quad (4.4-13)$$

Proof: The identity

$$(I + A + A^2 + \dots + A^n)(I - A) = I - A^{n+1} \quad (4.4-14)$$

and the fact that $|I-A| \neq 0$, imply that the inverse $(I-A)^{-1}$ exists. Thus

$$(I + A + A^2 + \dots + A^n) = (I - A)^{-1} - A^{n+1}(I - A)^{-1}.$$

(4.4-15)

Since $A^n \rightarrow 0$ as $n \rightarrow \infty$,

$$(I - A)^{-1} - A^{n+1}(I - A)^{-1} \rightarrow (I - A)^{-1} \text{ as } n \rightarrow \infty,$$

(4.4-16)

or equivalently

$$I + A + A^2 + \dots = (I - A)^{-1} . \quad (4.4-17)$$

Q.E.D.

The following theorem is stated in passing to show the accuracy of the n^{th} order series as an approximation for the matrix $(I - A)^{-1}$.

Theorem 6

If $|I-A| \neq 0$ and $A^n \rightarrow 0$ as $n \rightarrow \infty$ then

$$\left\| (I-A)^{-1} - (I+A+A^2+\dots+A^n) \right\| \leq \frac{\|A\|^{n+1}}{1 - \|A\|} . \quad (4.4-18)$$

Thus if we are able to show that the sequence A, A^2, A^3, \dots , tends to 0 and $|I-A| \neq 0$ then we can express $(I-A)^{-1}$ as the infinite series

$$\sum_{i=1}^{\infty} A^i = (I - A)^{-1} . \quad (4.4-19)$$

Also we can use an n^{th} order approximation for $(I - A)^{-1}$,

$$\sum_{i=1}^n A^i \approx (I - A)^{-1}, \quad (4.4-20)$$

and determine a measure of the error involved in (4.4-20). It is with this information that we will be able to derive another estimator for \underline{x} .

4.5 A First Order Estimator

Since a complete analytic solution is impractical, an approximation can be attempted using the theorems of the preceding section. Again the system of stochastic linear equations will be of the form

$$T \underline{x} = \underline{b}, \quad (4.5-1)$$

where T is an m by m matrix of random variables, \underline{b} is an m -dimensional vector of random variables and all of these random variables are independent. We wish to find either the distribution or the moments of \underline{x} . This system of equations may be rewritten as

$$(M_O + A) \underline{x} = \underline{b} \quad (4.5-2)$$

where $E[T] = M_O$ and hence $A = T - E[T]$. If the inverse of M_O exists then (4.5-2) can be written as

$$(I + A M_O^{-1}) M_O \underline{x} = \underline{b}, \quad (4.5-3)$$

where I is an m by m identity matrix. If the determinant of $I + A M_O^{-1}$ is non-zero, then its inverse exists and we have

$$M_O \underline{x} = (I + A M_O^{-1})^{-1} \underline{b} . \quad (4.5-4)$$

The probability that the matrix $I + A M_O^{-1}$ is singular is zero when the elements a_{ij} are continuous variates.

However, the probability of the matrix being singular may become non-zero if the determinant of the expected value of the matrix is close to zero and its variance is "large". Notice that $|E [I + A M_O^{-1}]| = 1$ so we need only to keep its variance "small" for $I + A M_O^{-1}$ to be non-singular.

We can now solve for \underline{x} in (4.5-4) since we have already assumed that M_O^{-1} exists, i.e.,

$$\underline{x} = M_O^{-1} (I + A M_O^{-1})^{-1} \underline{b} . \quad (4.5-5)$$

If we can show that $(A M_O^{-1})^n \rightarrow 0$ as $n \rightarrow \infty$, then we can apply Theorem 5 to $I + A M_O^{-1}$ since the $\Pr\{|I + A M_O^{-1}| \neq 0\}$ is one. From Theorem 4 we know that if $\|(A M_O^{-1})\| \leq 1$ then $(A M_O^{-1})^n \rightarrow 0$ as $n \rightarrow \infty$. If the $\Pr\{\|(A M_O^{-1})\| \leq 1\}$ approaches one (see Section 4.7) then it is reasonable to apply Theorem 5. Thus by Theorem 5, (4.5-5) becomes

$$\underline{x} = M_O^{-1} (I - A M_O^{-1} + (A M_O^{-1})^2 - \dots) \underline{b} \quad (4.5-6)$$

or equivalently

$$\underline{x} = (I - M_O^{-1} A + (M_O^{-1} A)^2 - \dots) M_O^{-1} \underline{b}. \quad (4.5-7)$$

For convenience we will write (4.5-7) as

$$\underline{x} = (I - M A + (M A)^2 - \dots) M \underline{b} \quad (4.5-8)$$

where $M = M_O^{-1}$.

From (4.5-8) it is now possible to define a first order estimator $\hat{\underline{x}}$ for \underline{x} as

$$\hat{\underline{x}} = (I - M A) M \underline{b} = M \underline{b} - M A M \underline{b}, \quad (4.5-9)$$

where the error $||\hat{\underline{x}} - \underline{x}||$ can be found by applying Theorem 6. This estimator is first order in A although it contains the cross product $M A M \underline{b}$. Next we will find the mean and the variance of this first order estimator.

Applying the expected value operator to (4.5-9), we find that

$$\begin{aligned} E[\hat{\underline{x}}] &= E[M \underline{b}] - E[M A M \underline{b}] \\ &= M E[\underline{b}] - M E[A] M E[\underline{b}]. \end{aligned} \quad (4.5-10)$$

Let $E[\underline{b}] = \underline{b}_O$ and recall that $E[A] = 0$. Thus the expected value of $\hat{\underline{x}}$ is

$$\therefore E[\hat{\underline{x}}] = M \underline{b}_O. \quad (4.5-11)$$

Unfortunately, the variance of \underline{x} is not as easy to derive. First, by definition

$$\text{Var}(\hat{\underline{x}} \hat{\underline{x}}') = E[\hat{\underline{x}} \hat{\underline{x}}'] - E[\hat{\underline{x}}] E'[\hat{\underline{x}}]. \quad (4.5-12)$$

Using (4.5-9), the matrix $\hat{\underline{x}} \hat{\underline{x}}'$ becomes

$$\hat{\underline{x}} \hat{\underline{x}}' = (M \underline{b} - M A M \underline{b}) (M \underline{b} - M A M \underline{b})' \quad (4.5-13)$$

or

$$\begin{aligned} \hat{\underline{x}} \hat{\underline{x}}' &= M \underline{b} \underline{b}' M' - M A M \underline{b} \underline{b}' M' - M \underline{b} \underline{b}' M' A' M' \\ &\quad + M A M \underline{b} \underline{b}' M' A' M'. \end{aligned} \quad (4.5-14)$$

The expected value operator is applied to (4.5-14) to obtain the first term of the variance (4.5-12), i.e.,

$$\begin{aligned} E[\hat{\underline{x}} \hat{\underline{x}}'] &= M E[\underline{b} \underline{b}'] M' - M E[A] M E[\underline{b} \underline{b}'] M' \\ &\quad - M E[\underline{b} \underline{b}'] M E[A'] M' + M E[A M \underline{b} \underline{b}' M' A'] M'. \end{aligned} \quad (4.5-15)$$

Recalling that $E[A] = 0$, we can simplify (4.5-15) and obtain

$$E[\hat{\underline{x}} \hat{\underline{x}}'] = M E[\underline{b} \underline{b}'] M' + M E[A M \underline{b} \underline{b}' M' A'] M'. \quad (4.5-16)$$

The expectation operator is an idempotent operator and thus (4.5-16) can be expressed as

$$E[\hat{\underline{x}} \hat{\underline{x}}'] = M E[\underline{b} \underline{b}'] M' + M E[A M E[\underline{b} \underline{b}'] M'A'] M'. \quad (4.5-17)$$

We could use (4.5-17) and (4.5-11) to express the variance (4.5-12), but this would be of little use. We need to express the $\text{Var}(\hat{\underline{x}} \hat{\underline{x}}')$ as a function of the known means and variances of A and \underline{b} in order to calculate the value of $\text{Var}(\hat{\underline{x}} \hat{\underline{x}}')$. To do this the following lemma is necessary.

Lemma: Let $A = [a_{ij}]$ be an m by m matrix of independent random variables such that

$$E[A] = 0 \quad (4.5-18)$$

and

$$\text{Var}(A) = S = [\sigma_{ij}^2]. \quad (4.5-19)$$

If $M = [m_{ij}]$ is an m by m matrix of constants, then

$$E[A M A'] = \mathcal{Q}(S \mathcal{Q}_{1,1}(M)). \quad (4.5-20)$$

\mathcal{Q} is a diagonalization operator which transforms the m -dimensional vector

$$S \mathcal{Q}_{1,1}(M) \quad (4.5-21)$$

into an m by m diagonal matrix. The operator $\mathcal{Q}_{1,1}$ ^(*) creates an m -dimensional vector from M such that

(*) The operator $\mathcal{Q}_{1,1}$ is a corruption of the APL dyadic operator \mathcal{Q} (see Falkoff and Iverson (1968) pp. 3.38-3.40).

$$\mathbb{Q}_{1,1}(M) = (m_{ii}) \quad , \quad i=1,2,\dots,m \quad . \quad (4.5-22)$$

Using (4.5-21) and (4.5-22) we can rewrite (4.5-20) as

$$E[A M A'] = \mathcal{D}(S \mathbb{Q}_{1,1}(M)) = [\delta_{ij} \sum_{j=1}^m \sigma_{ij}^2 m_{jj}] \quad , \\ i,j=1,2,\dots,m. \quad (4.5-23)$$

Proof: Letting $A' = P = [p_{ij}]$ and using the notation introduced in Chapter 2, we find

$$(M A')_{ij} = \sum_{k=1}^m m_{ik} p_{kj} = \sum_{k=1}^m m_{ik} a_{jk} . \quad (4.5-24)$$

Next the product $A M A'$ is such that

$$(A M A')_{ij} = \sum_{\ell=1}^m a_{i\ell} (M A')_{\ell j} = \sum_{\ell=1}^m \sum_{k=1}^m a_{i\ell} a_{jk} m_{\ell k} . \quad (4.5-25)$$

Now we take the expected value of $A M A'$ and consider the off-diagonal elements of this matrix, i.e.,

$$E[(A M A')_{ij}] = E\left[\sum_{\ell=1}^m \sum_{k=1}^m a_{i\ell} a_{jk} m_{\ell k}\right] , \quad i \neq j \quad (4.5-26)$$

or equivalently

$$E[(A M A')_{ij}] = \sum_{\ell=1}^m \sum_{k=1}^m m_{\ell k} E[a_{i\ell} a_{jk}] = 0 , \quad i \neq j . \quad (4.5-27)$$

The expected value of the diagonal element of $A M A'$ is

$$E[(A M A')_{ii}] = E\left[\sum_{\ell=1}^m a_{i\ell}^2 m_{\ell\ell} + \sum_{\ell=1}^m \sum_{\substack{k=1 \\ k \neq \ell}}^m a_{i\ell} a_{\ell k} m_{\ell k}\right] \quad (4.5-28)$$

or

$$E[(A M A')_{ii}] = \sum_{\ell=1}^m m_{\ell\ell} E[a_{i\ell}^2] + \sum_{\ell=1}^m \sum_{\substack{k=1 \\ k \neq \ell}}^m m_{\ell k} E[a_{i\ell} a_{\ell k}]. \quad (4.5-29)$$

Since a_{ij} and $a_{\ell k}$ for $i \neq \ell$ and $j \neq k$ are independent (4.5-29) can be rewritten as

$$E[(A M A')_{ii}] = \sum_{\ell=1}^m m_{\ell\ell} \sigma_{i\ell}^2, \quad i=1, 2, \dots, m. \quad (4.5-30)$$

The conclusion (4.5-23) is derived from (4.5-27) and (4.5-30).

Q.E.D.

For the vector $\underline{b} = (b_i)$ let the $E[\underline{b}] = \underline{b}_o$ and $\text{var}(\underline{b}) = \underline{\lambda} = (\lambda_i)$, and therefore

$$\text{Var}(\underline{b} \underline{b}') = E[\underline{b} \underline{b}'] - E[\underline{b}] E'[\underline{b}] \quad (4.5-31)$$

or

$$\mathcal{D}(\underline{\lambda}) = E[\underline{b} \underline{b}'] - \underline{b}_o \underline{b}'_o, \quad (4.5-32)$$

where $\mathcal{D}(\underline{\lambda})$ is a diagonal matrix of the form

$$\mathcal{D}(\underline{\lambda}) = [\delta_{ij} \lambda_i^2] . \quad (4.5-33)$$

We now return to the problem of expressing $E[\hat{\underline{x}} \hat{\underline{x}}']$ as a function of the means and variances of \underline{A} and \underline{b} .

Using (4.5-32) and (4.5-17), we obtain

$$E[\hat{\underline{x}} \hat{\underline{x}}'] = M E[\underline{b} \underline{b}'] M' + M E[A M (\mathcal{D}(\underline{\lambda}) + \underline{b}_O \underline{b}_O') M' A'] M' \quad (4.5-34)$$

or

$$\begin{aligned} E[\hat{\underline{x}} \hat{\underline{x}}'] &= M E[\underline{b} \underline{b}'] M' + M E[A M \mathcal{D}(\underline{\lambda}) M' A'] M' \\ &\quad + M E[A M \underline{b}_O \underline{b}_O' M' A'] M' . \end{aligned} \quad (4.5-35)$$

We can apply the preceding lemma to $U = M \mathcal{D}(\underline{\lambda}) M'$ and $V = M \underline{b}_O \underline{b}_O' M'$ and obtain

$$E[\hat{\underline{x}} \hat{\underline{x}}] = M \{E[\underline{b} \underline{b}'] + \mathcal{D}(S Q_{1,1}(U)) + \mathcal{D}(S Q_{1,1}(V))\} M'. \quad (4.5-36)$$

From (4.5-36) and (4.5-11) the variance-covariance matrix (4.5-12) becomes

$$\begin{aligned} \text{Var}(\hat{\underline{x}} \hat{\underline{x}}) &= M \{E[\underline{b} \underline{b}'] + \mathcal{D}(S Q_{1,1}(U)) + \mathcal{D}(S Q_{1,1}(V))\} M' \\ &\quad - M \underline{b}_O \underline{b}_O' M' \end{aligned} \quad (4.5-37)$$

or

$$\begin{aligned} \text{Var}(\hat{\underline{x}} \hat{\underline{x}}') &= M \{E[\underline{b} \underline{b}'] - \underline{b}_O \underline{b}_O' + \mathcal{D}(S Q_{1,1}(U)) \\ &\quad + \mathcal{D}(S Q_{1,1}(V))\} M' . \end{aligned} \quad (4.5-38)$$

Finally, if the identity (4.5-32) is used in (4.5-38) we obtain the result

$$\text{Var}(\hat{\underline{x}} \hat{\underline{x}}') = M \{ \mathcal{D}(\underline{\lambda}) + \mathcal{D}(S Q_{1,1}(U)) + \mathcal{D}(S Q_{1,1}(V)) \} M' , \quad (4.5.39)$$

where

$$\mathcal{D}(\underline{\lambda}) = [\delta_{ij} \lambda_i^2] , \quad (4.5-40)$$

$$\mathcal{D}(S Q_{1,1}(U)) = [\delta_{ij} \sum_{j=1}^m \sigma_{ij}^2 (U)_{jj}] , \quad (4.5-41)$$

$$U = M \mathcal{D}(\underline{\lambda}) M' ,$$

and

$$\mathcal{D}(S Q_{1,1}(V)) = [\delta_{ij} \sum_{j=1}^m \sigma_{ij}^2 (V)_{jj}] , \quad (4.5-42)$$

$$V = M \underline{b}_o \underline{b}_o' M' .$$

In conclusion, we now have an expression (4.5-39) for the variance-covariance matrix of $\hat{\underline{x}}$ which is stated in terms of the known means and variances of A and \underline{b} . Also the formula for the $\text{Var}(\hat{\underline{x}} \hat{\underline{x}}')$ is fairly simple involving only matrix algebra.

We will now consider the problem discussed by Tintner (1955) in which only T is random and its elements are independent normal variates with mean M_o and variance matrix $S = [\sigma_{ij}^2]$. The first order

estimator $\hat{\underline{x}}$ for this problem is

$$\hat{\underline{x}} = M \underline{b} - M A M \underline{b} , \quad (4.5-43)$$

where $A = T - M_O^{-1}$ and $M = M_O^{-1}$, and $\hat{\underline{x}}$ is a linear combination of the normal variates in A . Thus if $\hat{\underline{x}} = (\hat{x}_j)$, then \hat{x}_j , $j=1, 2, \dots, m$ is normally distributed with

$$E [\hat{x}_j] = \sum_{k=1}^m m_{jk} b_k , \quad j=1, 2, \dots, m , \quad (4.5-44)$$

and

$$\text{Var}(\hat{x}_j) = \sum_{i=1}^m \sum_{\ell=1}^m \sum_{k=1}^m m_{j\ell}^2 \sigma_{\ell i}^2 m_{ik}^2 b_k^2 , \quad j=1, 2, \dots, m . \quad (4.5-45)$$

We would obtain the same result if we used Tintner's estimator in place of the first order estimator.

If as in Tintner and Raghavan (1965), we assume that only \underline{b} is random and that $T = M_O$, then the stochastic system may be written as

$$M_O \underline{x} = \underline{b} , \quad \text{i.e., } A = 0 . \quad (4.5-46)$$

Thus the first order estimator, $\hat{\underline{x}}$, is

$$\hat{\underline{x}} = M_O^{-1} \underline{b} - M_O^{-1} 0 M_O^{-1} \underline{b} = M_O^{-1} \underline{b} \quad (4.5-47)$$

which is the same as the vector \underline{x} in (4.5-46). If b_i , where $\underline{b} = (b_i)$, is normally distributed then x_j is a linear combination of normal variates, and therefore x_j is a normal variate with

$$E[x_j] = \sum_{k=1}^m m_{jk}^2 E[b_k] \quad (4.5-48)$$

and

$$\text{Var}(x_j) = \sum_{k=1}^m m_{jk}^2 \lambda_k^2 . \quad (4.5-49)$$

Again if we started with (4.3-5) instead of (4.5-43) then we could still derive the same results (4.5-48) and (4.5-49).

The only difference between Tintner's approximation (4.3-5) and the first order approximation developed here occurs when both A and \underline{b} are random variables. If we accept the arguments which lead to derivation of \hat{x} , then it does not make sense to use \hat{x}_t except when either A or \underline{b} are random. However, if we still insist on using \hat{x}_t for convenience, then of necessity we must neglect the combined influence of A and \underline{b} since \hat{x}_t is merely a linear combination of A and \underline{b} . Whereas \hat{x} includes a cross product term of A and \underline{b} . Notice that there is no reason why a higher order estimator could not be developed from (4.5-8) and a similar analysis applied to obtain its mean vector and variance-covariance matrix.

4.6 Simulation of Stochastic Linear Equations

We have now derived an estimator,

$$\hat{x} = M \underline{b} - M A M \underline{b} , \quad (4.6-1)$$

for the solution \underline{x} of the system of stochastic linear equations

$$(M_0 + A) \underline{x} = \underline{b}, \quad (4.6-2)$$

with known probability density function $P(A, \underline{b})$. It would be ideal to test this estimator on a "real world" problem. Unfortunately, because of the lack of familiarity with such stochastic models very few examples of them have been explored. This does not infer that such problems do not exist, but as was pointed out earlier, the solution of a stochastic model is very complicated and most persons would prefer to ignore the stochastic nature of the problem and solve the deterministic model. Thus to test the approximation (4.6-1) and to analyse the stochastic model (4.6-2), a stochastic model was simulated.

There are two methods of simulating the model (4.6-2). The Monte Carlo method could be used to generate problems of the form (4.6-2), to solve these problems, and to collect the data on the system's solution. We can use the data collected from these experiments to predict the behaviour of the model. The problem generation is done by selecting randomly with the aid of a pseudo-random number generator, a point from the population with the given probability

density function of A and \underline{b} . The pseudo-random number generator is described in Appendix 1, along with the documentation of the random number tests. The second method of simulation is the method used in most of Tintner's work (1955, 1965) and involves approximating the given probability density function with a step function, i.e., the continuous density is replaced by a discrete density,

$$0 \leq \Pr \{ \underline{\eta} = (a_{11}, \dots, a_{mm}, b_1, \dots, b_m) | \underline{\eta} \in S \subset \mathbb{R}^{m^2+m} \}, \quad (4.6-3)$$

where the value set $S = \{(A_i, b_i) | i=1,2,\dots,k\}$ is used to generate all possible problems. This method may be viewed as a sampling procedure in \mathbb{R}^{m^2+m} , where certain points are selected and assigned a specific weight. We will refer to this method of simulation as "Sampling".

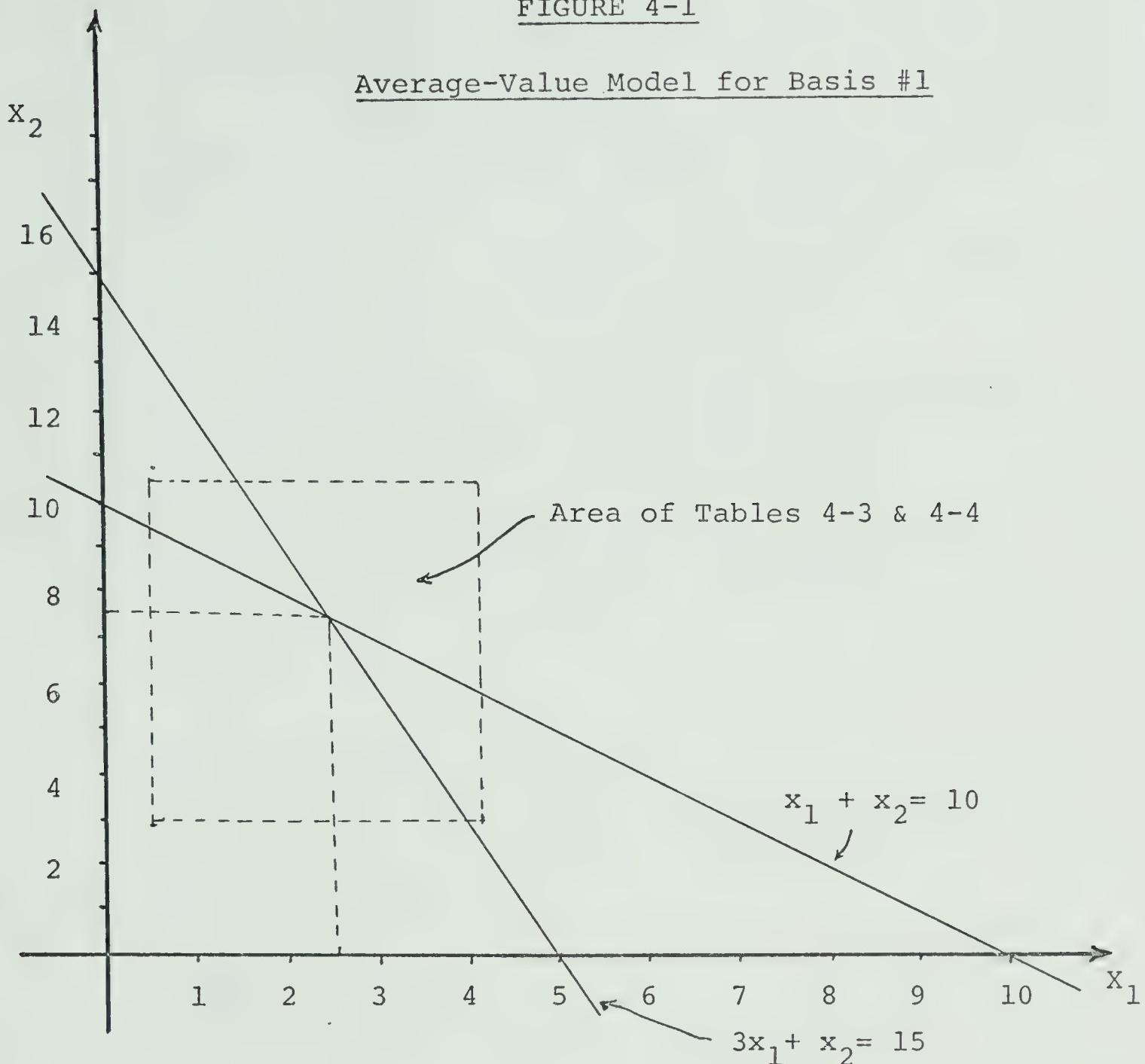
A Monte Carlo simulation was performed on the following example.

Basis #1

Solve

$$(M_0 + A) \underline{x} = \underline{b}, \quad (4.6-4)$$

where A is a 2 by 2 matrix of independent normal variates with mean zero and variance

FIGURE 4-1Average-Value Model for Basis #1

Solve:

$$3x_1 + x_2 = 15$$

$$x_1 + x_2 = 10$$

$$x_1 = 2.5, \quad x_2 = 7.5$$

$$\text{Var}(A) = \begin{bmatrix} 0.04 & 0.01 \\ 0.09 & 0.0016 \end{bmatrix} . \quad (4.6-5)$$

The matrix M_o is a 2 by 2 matrix of constants,

$$M_o = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} , \quad (4.6-6)$$

and \underline{b} is a 2-dimensional vector of independent normal variates with mean $\underline{b}_o = (15, 10)$ and variance $\text{Var}(b) = (0.25, 0.36)$. The components of A are independent of the components of \underline{b} . The average value model,

$$M_o \underline{x} = \underline{b}_o , \quad (4.6-7)$$

is depicted in Figure 4-1.

If we use the first order estimator for Basis #1 we find the mean of $\hat{\underline{x}}$ is

$$E[\hat{\underline{x}}] = \begin{bmatrix} 2.5 \\ 7.5 \end{bmatrix} \quad (4.6-8)$$

and the variance-covariance matrix of $\hat{\underline{x}}$ is

$$\text{Var}(\hat{\underline{x}} \hat{\underline{x}}') = \begin{bmatrix} 0.52624 & -1.04005 \\ -1.04005 & 2.58149 \end{bmatrix} . \quad (4.6-9)$$

To simulate this model it was necessary to generate a matrix A^* and a vector \underline{b}^* from the distributions defined above. This involves the generation of points from a

normal distribution which was accomplished by using the Box-Muller algorithm (Box and Muller (1958)). After each problem is generated, \underline{x} is found, and the data are collected for the mean, variance and covariance of \underline{x} . A two-way frequency table for \underline{x} was also maintained. As well as solving and accumulating data for \underline{x} , we also recorded the values of

$$\hat{\underline{x}}_i = \{I + \sum_{j=1}^i (-1)^j (M A^*)^j\} M \underline{b}^* \quad (4.6-10)$$

for $i=1, 2, 3, 4, 8, 16$. This model was then simulated for 1000 systems of random equations. The means, variances, and covariances are shown in Table 4-1.

TABLE 4-1
Results for Basis #1 Using 1000 Random Equations

$\hat{\underline{x}}_i$	Mean		Variance		$Cov(x_1, x_2)$
	x_1	x_2	x_1	x_2	
*	2.50000	7.50000	0.52624	2.58149	-1.04005
1	2.48642	7.54649	0.53133	2.57098	-1.04123
2	2.53605	7.40882	0.56440	2.75671	-1.11525
3	2.53521	7.40875	0.59563	2.94980	-1.19093
4	2.54070	7.39570	0.60148	2.98074	-1.20420
8	2.54185	7.39285	0.60957	3.02917	-1.22381
16	2.54195	7.39261	0.61015	3.03298	-1.22529
0	2.54196	7.39261	0.61015	3.03301	-1.22531

* Theoretical Values of $\hat{\underline{x}}_1$.

Relative Frequency of $\{|M_O^{-1} A| < 1\}$ is 1.00000

$Pr(|M_O^{-1} A| < 1) \geq 0.90600 = s$

The approximation 0 in Table 4-1 is the variable \underline{x} while the approximation * is the theoretical values for $\hat{\underline{x}}_1$ calculated from the formulae of the previous section. Notice that the moments of the simulated first order estimator agree with the theoretical values. We can also see the convergence of the moments of the i^{th} order estimator to the moments of \underline{x} as i increases. The last two entries in the table will be explained in the next section.

Table 4-2 contains the results of generating 20000 systems of random equations for Basis #1. Only \underline{x} and $\hat{\underline{x}}_1$ were generated because of the time required.

TABLE 4-2

Results For Basis #1 Using 20000 Random Equations

$\hat{\underline{x}}_i$	Mean		Variance		$\text{Cov}(\underline{x}_1, \underline{x}_2)$
	\underline{x}_1	\underline{x}_2	\underline{x}_1	\underline{x}_2	
*	2.50000	7.50000	0.52624	2.58148	-1.04005
1	2.49875	7.49599	0.52650	2.59586	-1.04464
0	2.56176	7.32651	0.63109	3.18826	-1.28731

* Theoretical Values for $\hat{\underline{x}}_1$.

Relative Frequency of $\{|M_O^{-1} A| < 1\}$ is 1.00000

$$\Pr(|M_O^{-1} A| < 1) = 0.90600 = s$$

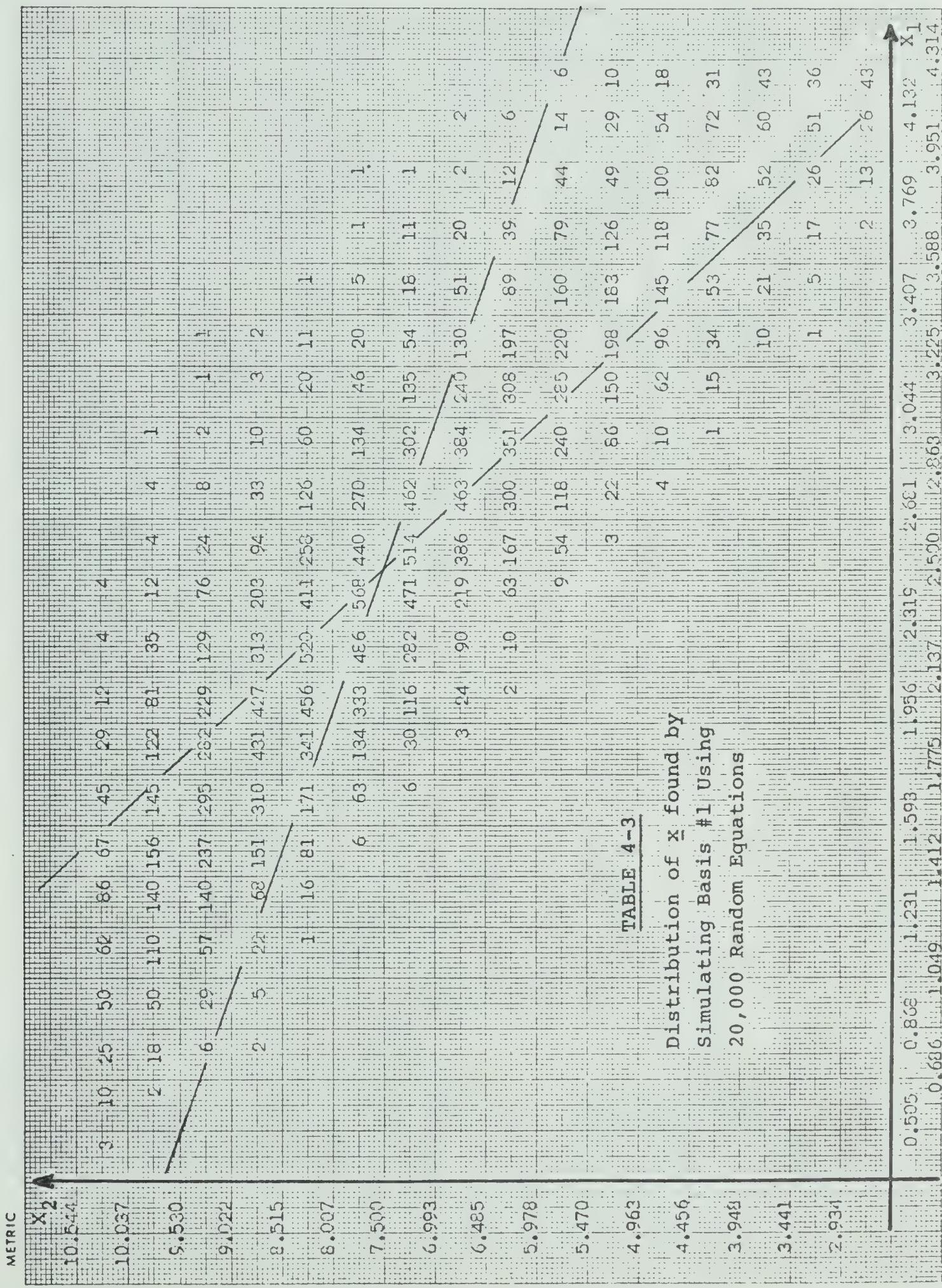


TABLE 4-3

Distribution of \underline{x} found by
Simulating Basis #1 Using
20,000 Random Equations

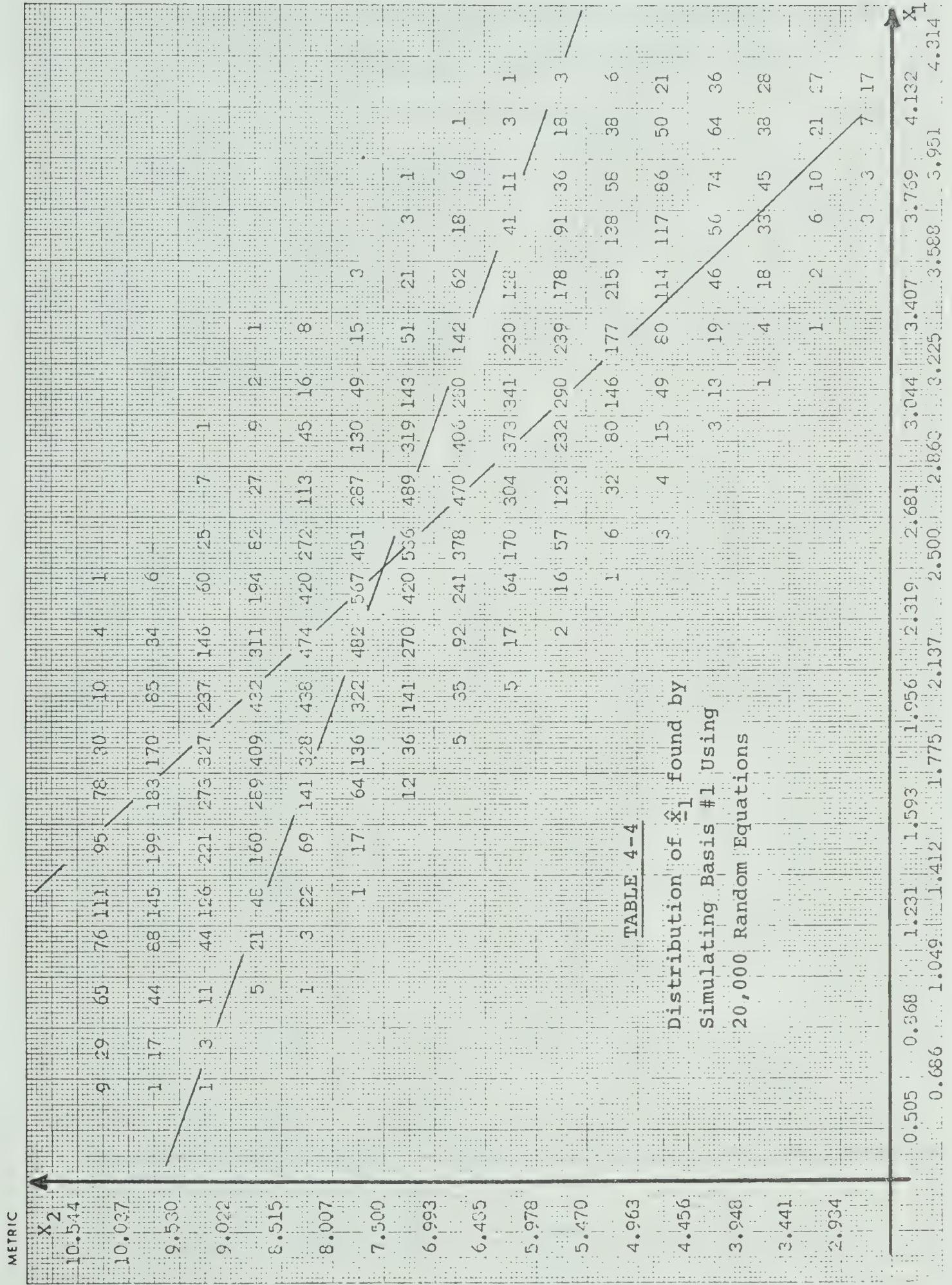


TABLE 4-4

Distribution of \hat{x}_1 found by Simulating Basis #1 Using 20,000 Random Equations

This experiment shows even better the agreement between the simulated moments and the theoretical moments of \hat{x}_1 . The two-way frequency tables for x and \hat{x}_1 , using 20000 systems of random equations are given in Tables 4-3 and 4-4, respectively. The number in each cell represents the number of times x or \hat{x}_1 (as the case may be) falls into that cell. The boundary points of each cell appear on the appropriate axes.

TABLE 4-5

Results For Basis #4 Using 1000 Random Equations

\hat{x}_1	Mean		Variance		$Cov(x_1, x_2)$
	x_1	x_2	x_1	x_2	
*	10.00000	5.00000	0.52058	1.77418	-0.52058
1	10.02329	4.94773	0.52244	1.94583	-0.55135
2	10.03957	4.93229	0.52448	1.95511	-0.55412
3	10.03959	4.9224	0.52606	1.95976	-0.55556
4	10.03966	4.93217	0.52609	1.95979	-0.55558
8	10.03966	4.93217	0.52607	1.95977	-0.55557
16	10.03966	4.93217	0.52607	1.95977	-0.55557
0	10.03967	4.93217	0.52610	1.95982	-0.55559

* Theoretical Values for \hat{x}_1

Relative Frequency of $\{|M_O^{-1} A| < 1\}$ is 1.00000

$$\Pr(|M_O^{-1} A| < 1) \geq 0.98840 = s$$

TABLE 4-6
Results For Basis #5 Using 1000 Random Equations

\hat{x}_i	Mean		Variance		$\text{Cov}(x_1, x_2)$
	x_1	x_2	x_1	x_2	
*	15.00000	-5.00000	2.50250	3.22290	-2.502550
1	14.95639	-4.93503	2.72544	3.49708	-2.74410
2	15.11855	-5.09578	2.69834	3.44607	-2.70560
3	15.11490	-5.09194	2.85679	3.60863	-2.86204
4	15.12008	-5.09705	2.85827	3.60697	-2.86204
8	15.12055	-5.09755	2.84957	3.59783	-2.85328
16	15.12056	-5.09755	2.84952	3.59778	-2.87323
0	15.12016	-5.09710	2.86757	3.61600	-2.87105

* Theoretical Values for \hat{x}_1

Relative Frequency of $\{|M_O^{-1} A| < 1\}$ is 1.00000

$$\Pr(|M_O^{-1} A| < 1) \geq 0.98840 = s$$

Tables 4-5 and 4-6 are similar to Table 4-1 except that they contain the corresponding results for the examples referred to as Basis #4 and Basis #5, respectively.

Basis #4

This example is also defined by the model (4.6-4) except that

$$\text{Var}(A) = \begin{bmatrix} 0.01 & 0.0 \\ 0.0016 & 0.0 \end{bmatrix} \quad (4.6-11)$$

and

$$M_O = \begin{bmatrix} 1.0 & 1.0 \\ 1.0 & 0.0 \end{bmatrix} . \quad (4.6-12)$$

Basis #5

The Basis #5 model is again the same as Basis #1 except that

$$\text{Var}(A) = \begin{bmatrix} 0.01 & 0.0 \\ 0.0016 & 0.0 \end{bmatrix} \quad (4.6-13)$$

and

$$M_O = \begin{bmatrix} 1.0 & 0.0 \\ 1.0 & 1.0 \end{bmatrix} . \quad (4.6-14)$$

Again these models show the trends pointed out earlier in Table 4-1 regarding the convergence of higher order estimators and the behaviour of the first order estimator. In summary, these results show that the first order estimator \hat{x}_1 is a reasonable approximation to \underline{x} and that the formulae for the moments of this estimator are correct. The results also indicate that a more accurate approximation could be given if a higher order estimator were used. However, as is obvious from the last section, the calculation of the theoretical moments of higher order estimator is very difficult. Thus the use of \hat{x}_1 as an estimator of \underline{x} is the most practical solution to the problem.

Tintner's Sampling technique (1955) was used for the Basis #1 model to compare its results with those of the Monte Carlo method. Discrete approximations were used for the standard normal distribution. These approximations involved the use of 3, 5, or 7 points with the corresponding probabilities and these approximations are shown in Table 4-7.

TABLE 4-7

Discrete Density Functions for Sampling Technique

3 Point		5 Point		7 Point	
X	Pr (X=x)	X	Pr (X=x)	X	Pr (X=x)
-1.5	0.22663	-2.0	0.06681	-2.25	0.03040
0.0	0.54675	-1.0	0.24173	-1.50	0.09990
+1.5	0.22663	0.0	0.38293	-.75	0.22354
		1.0	0.24173	0.0	0.29234
		2.0	0.06681	0.75	0.22354
				1.50	0.09990
				2.25	0.03040

The approximations are translated to fit the appropriate normal distributions. Note that if there are k possible values for each of the $m^2 + m$ variables (A and b) then there are k^{m^2+m} possible problems to be solved, i.e.,

729, 15625, and 117649 problems for the 3, 5, and 7 point approximations, respectively. Thus it is apparent that the number of possible problems is large even when the values of m and k are small. Also the results of the sampling experiments are biased by the particular approximation used, i.e., increasing the number of points sampled does not necessarily improve the results of the simulation. A comparison of the results for \underline{x} , generated by the Monte Carlo and the Sampling techniques is shown in Table 4-8.

TABLE 4-8
Comparison of Monte Carlo and Sampling

Statistics for \underline{x}	Number of Equations Solved				
	Monte Carlo		Sampling Technique		
	1000	20000	729	15625	117649
mean x_1	2.54166	2.55978	2.55978	2.55738	2.53344
mean x_2	7.39369	7.30891	7.33496	7.32792	7.28589
var(x_1)	0.60690	0.61953	0.60150	0.62528	0.71909
var(x_2)	3.01292	3.37146	2.99744	3.05647	3.02632
cov(x_1, x_2)	-1.21769	-1.24744	-1.20839	-1.21999	-1.06708

4.7 A Necessary Condition For The First Order Estimator

As was pointed out in Section 4.5 a necessary condition for the existence of the first order estimator \hat{x} is that

$$\Pr(||A M|| \leq 1) \text{ is near 1} . \quad (4.7-1)$$

Recall that A is a matrix of independent random variables and $M = M_O^{-1}$ is a matrix of constants. Thus for the condition (4.7-1) to be useful we must first define the specific norm under consideration. The norm in (4.7-1) will be defined to be the Euclidean norm, i.e.,

$$||T|| = \left(\sum_{i=1}^m \sum_{j=1}^m t_{ij}^2 \right)^{1/2} , \quad (4.7-2)$$

or equivalently

$$||T||^2 = \sum_{i=1}^m \sum_{j=1}^m t_{ij}^2 . \quad (4.7-3)$$

Notice the Euclidean norm satisfies the conditions (4.4-6) through (4.4-10). Secondly, we need the following theorem which is stated here without proof (see Hogg and Craig (1968) p.47):

Theorem 7

Let $u(X)$ be a nonnegative function of the random variable X . If $E[u(X)]$ exists, then, for every positive constant c ,

$$\Pr(u(X) \geq c) \leq \frac{E[u(X)]}{c} . \quad (4.7-4)$$

Clearly $\|A_M\|^2$ is a nonnegative function of the random variable A_M , and

$$(A_M)_{ij} = \sum_{k=1}^m a_{ik} m_{kj} . \quad (4.7-5)$$

Thus substituting the formula (4.7-5) into the definition of the norm, we obtain

$$\|A_M\|^2 = \sum_{i=1}^m \sum_{j=1}^m \left(\sum_{k=1}^m a_{ik} m_{kj} \right)^2 \quad (4.7-6)$$

or

$$\|A_M\|^2 = \sum_{i=1}^m \sum_{j=1}^m \left\{ \sum_{k=1}^m a_{ik}^2 m_{kj}^2 + 2 \sum_{\substack{r=1 \\ r \neq s}}^m \sum_{\substack{s=1 \\ s \neq r}}^m a_{ir} a_{is} m_{rj} m_{sj} \right\} . \quad (4.7-7)$$

Taking the expected value of (4.7-7), we have

$$E[\|A_M\|^2] = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m m_{kj}^2 \sigma_{ik}^2 , \quad (4.7-8)$$

since $E[a_{ij}] = 0$; $\text{Var}(a_{ij}) = \sigma_{ij}^2$, and $E[a_{ij} a_{kl}] = 0$ if $i \neq k$ and $j \neq l$.

The condition that the $\Pr(\|A_M\| \leq 1)$ be near one is equivalent to saying that $\Pr(\|A_M\| > 1)$ be near zero. For this latter probability we can apply Theorem 7 with $c=1$ and $u(X) = \|A_M\|^2$ and obtain

$$\Pr(\|A M\|^2 > 1) \leq E[\|A M\|^2] . \quad (4.7-9)$$

This inequality can be rewritten using (4.7-8) to obtain

$$\Pr(\|A M\| > 1) < \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m m_{kj}^2 \sigma_{ik}^2 . \quad (4.7.10)$$

We can again rewrite this last relation in the form

$$\Pr(\|A M\| < 1) \geq 1 - \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m m_{kj}^2 \sigma_{ik}^2 . \quad (4.7-11)$$

Therefore, if the right-hand side of (4.7-11) is near one then it makes sense to use the first order estimator.

The usefulness of the measure

$$s = 1 - \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m m_{kj}^2 \sigma_{ik}^2 , \quad (4.7-12)$$

can be seen by comparing the rate of convergence of the higher order estimators \hat{x}_j to \bar{x} . In comparing Table 4-1 with Table 4-9, we see that the value of s for Basis #1 is 0.90600 and the convergence is very good, while the value of s for Basis #2 is 0.87000 and the convergence is not as good. Notice that the relative frequency of the event $\|A M\| < 1$ is shown at the bottom of Tables 4-1, 4-5, 4-6, and 4-9. Notice also that this relative frequency is always greater than the corresponding value of s , as is indicated by the relation (4.7-11).

TABLE 4-9Results for Basis #2 Using 1000 Random Equations

\hat{x}_i	Mean		Variance		$\text{Cov}(x_1, x_2)$
	x_1	x_2	x_1	x_2	
*	10.00000	-15.00000	9.39240	88.79600	-28.17720
1	10.01799	-15.12056	9.68515	91.58667	-29.01191
2	10.93386	-17.85553	11.47379	107.97150	-34.34419
3	10.94053	-17.88035	17.52639	162.26512	-52.33118
4	11.17842	-18.58759	20.10144	184.68828	-59.89330
8	11.30922	-18.97802	19.32980	178.17596	-57.67543
16	11.37631	-19.18282	22.06726	204.61832	-66.18304
0	12.42601	-22.26652	1064.38	9322.40	too large
Relative Frequency of $\{ A_M < 1 \} = 1.00000$ $\Pr(A_M < 1) \geq 0.87000 = s$					

CHAPTER 5

STOCHASTIC LINEAR PROGRAMMING

5.1 Introduction

We saw in Chapter 4 that a stochastic linear program is simply a classical linear program in which some elements are random variables. A stochastic linear programming model may be defined by (4.1-1), (4.1-2), (4.1-3), and (4.1-4). The solution to this model will be the specification of the distribution of the objective function. Stochastic linear programming may also be used to analyse the implications of the stochastic model as it is related to decision-making under risk (see Section 5.4). We found in the previous chapter that very little progress has been made in the analytic solution of the system of stochastic linear equations, and thus there has been even less success in the analytic solution of the stochastic linear program. Tintner (1955) suggests an analytic approach to the solution of a stochastic linear program, and this will be the topic of the next section. However, as was found for a system of stochastic linear equations, the solution of a stochastic linear program is obtained by a numerical approximation, i.e., an approximate solution is found by simulating (or sampling) the stochastic program.

The stochastic linear programming model has been applied in several areas of economic planning. Tintner

(1960) used this model to analyse agricultural production, and Tintner and Raghavan (1965) used this model in a dynamic planning model for India. Charnes and Cooper (1964) also applied stochastic programming to critical path scheduling in which the task completion times were random.

5.2 Analytic Solution

The only attempt at an analytic solution to a stochastic linear program was presented by Tintner (1955). Unfortunately, the only success his method has had is in solving a single-constraint model with two decision variates. This lack of success is due to the method's complexity. Hence the discussion of this method is included for completeness, rather than as a suggested practical solution.

Let us define the stochastic linear program as

$$z^* = \max_{\underline{x}} z = \max_{\underline{x}} \underline{c}' \underline{x} \quad (5.2-1)$$

subject to

$$A \underline{x} = \underline{b} \quad (5.2-2)$$

and

$$\underline{x} \geq \underline{0} , \quad (5.2-3)$$

where A is an m by n+m matrix, and \underline{b} , \underline{c} , and \underline{x} are

dimensioned accordingly. This model is the same as the model introduced in Section 4.1 except that slack variables have been added. The elements A , \underline{b} , and \underline{c} may be random, with the known joint probability density function

$$\Pr((A, \underline{b}, \underline{c}) = \underline{\eta} \mid \underline{\eta} \in S) = R e^{(m+n)m+2m+n} . \quad (5.2-4)$$

The set S will be referred to as the parameter space.

For such a stochastic linear programming problem, there are $K = \binom{m+n}{m}$ possible bases $\{\mathcal{B}^{(k)}\}$. If we define $\underline{x}^{(k)}$, $k=1,2,\dots,K$ as the decision vector corresponding to the k^{th} basis, then we can define the set $S_k \subset S$ as the region in which the k^{th} selection is feasible, i.e.,

$$S_k = \{ \underline{\eta} \mid \underline{x}^{(k)} \geq \underline{0} \text{ and } \underline{\eta} \in S \} . \quad (5.2-5)$$

Tintner claims that the regions S_1, S_2, \dots, S_K form a system of overlapping subregions of S . Notice that for any given triple $(A, \underline{b}, \underline{c}) = \underline{\eta}$, there exists an r such that

$$z^{(r)} = \max_k z^{(k)}, \quad k=1,2,\dots,K. \quad (5.2-6)$$

We can then define a region T_r such that

$$T_r = \{ \underline{\eta} \mid z^{(r)} = \max_k z^{(k)} \text{ and } \underline{\eta} \in S \} , \quad (5.2-7)$$

or T_r is the region in which the r^{th} selection is maximal. Tintner claims that the regions T_k , $k=1,2,\dots,K$, do not overlap and thus the regions $S_k \cap T_k = U_k$, $k=1,2,\dots,K$ do not overlap. If $\underline{\eta} \in U_k \subset S$, then the k^{th} selection is both feasible and optimal.

Stochastic linear programming maps the set S into R^e where this mapping may be a many-to-one mapping.

However, $S = \bigcap_{k=1}^K U_k$ where U_k , $k=1,2,\dots,K$, are mutually disjoint, and the mapping from U_k to R^e is one-to-one.

The transformation from U_k to R^e is the same problem discussed in Chapter 4 where

$$\underline{x}^{(k)} = (\mathcal{B}^{(k)})^{-1} \underline{b} \geq 0 \quad (5.2-8)$$

and

$$z^{(k)} = (\underline{c}^{(k)})' \underline{x}^{(k)} . \quad (5.2-9)$$

We have shown how to determine the probability density function of $\underline{x}^{(k)}$, and from this density, it is possible to derive the density of $z^{(k)}$. Thus we can derive K one-to-one mappings of the mutually disjoint, exhaustive subsets U_k of S into R^e . This again permits us to use the transformation of variables technique to obtain the distribution of z^* and S (see Hogg and Craig (1965) pp. 131, 132).

If the reader thought the transformations of the previous chapter were complicated, he is advised to study Tintner's example ((1955) pp. 205-15) of this technique for $(m, n) = (1, 2)$. Again it should be stressed that while the above arguments may lead to an analytic solution, their complexity eliminates any value they might have as a practical procedure. Stochastic programming can best be solved by using an approximation obtained by simulating the model.

However, before proceeding with a discussion on simulation, we will digress to present an approximation which has not been mentioned in the literature. It is possible to approximate the distribution $z^{(k)}$ for a particular basis and to calculate the mean and variance of $z^{(k)}$. Why not determine the expected value of $E[z^{(k)}]$ and $\text{Var}(z^{(k)})$ over all possible bases? That is, if the probability that the k^{th} basis is optimal is p_k , then

$$\begin{aligned}\bar{\mu}_z &= E_k [E[z^{(k)}]] = \sum_{k=1}^K p_k E[z^{(k)}] \\ &= \sum_{k=1}^K \sum_{i=1}^m p_k c_i^{(k)} E[x_i^{(k)}]\end{aligned}\quad (5.2-10)$$

and

$$\begin{aligned}\bar{\sigma}_z^2 &= E_k [\text{Var}(z^{(k)})] = \sum_{k=1}^K p_k \text{Var}(z^{(k)}) \\ &= \sum_{k=1}^K p_k \left\{ \sum_{i=1}^m (c_i^{(k)})^2 \text{Var}(x_i^{(k)}) \right. \\ &\quad \left. + 2 \sum_{i < j} \sum c_i^{(k)} c_j^{(k)} \text{Cov}(x_i^{(k)}, x_j^{(k)}) \right\}.\end{aligned}\quad (5.2-11)$$

The probability that a particular basis is optimal is unknown and is still an unsolved problem. However, to illustrate the above method, we will set p_k equal to the relative frequency of the k^{th} basis being optimal. We will use the figures obtained from simulating 19000 random linear programs corresponding to Basis #1 (i.e., the equality of (4.6-4) is replaced by the inequality less-than-or-equal-to). The values of p_k , $k=1, 4$, and 5 are taken to be the relative frequencies of the pairs (x_1, x_2) , (x_2, x_3) , and (x_2, x_4) , respectively, obtained from Table 5-2. The values of $\underline{x}^{(k)}$, $k=1, 4$, and 5 can be taken from (row 0) Tables 4-1, 4-5, and 4-6, respectively. The relative frequencies of the bases $k=2, 3$, and 6 are zero. Using these figures, we can calculate the approximate value of (5.2-10) and (5.2-11). These values are contained in Table 5-1.

TABLE 5-1
Weighted Bases Technique

Method	$\bar{\mu}_z$	$\bar{\sigma}_z^2$
$\underline{x}^{(k)}$	19.94817	2.38455
$\hat{x}^{(k)}$	19.861	2.30965
Simulation	20.05181	2.08087

We could adapt this method by employing the first-order estimator $\hat{\underline{x}}^{(k)}$ in place of $\underline{x}^{(k)}$. The last row of Table 5-1 shows the results of simulating 19000 linear programs. We can see that this method provides a crude estimate of the moments of z^* .

5.3 Simulation of Stochastic Programs

The approximate distribution of z may be found for

$$z = \max_{\underline{x}} \underline{c}' \underline{x} \quad (5.3-1)$$

subject to

$$(\underline{M}_0 + A) \underline{x} \leq \underline{b} \quad (5.3-2)$$

and

$$\underline{x} \geq \underline{0}$$

by simulating such problems according to the given probability density function. This procedure is similar to the method employed in Section 4.6. We select at random points from a population defined by (5.2-4), generate the corresponding linear programs, and then solve these programs retaining the pertinent data. Again this simulation technique is referred to as the Monte Carlo technique while the Sampling method refers to sample points and their weights. The methodological details of these simulations can be found in Appendix 2.

To analyse the behaviour of the stochastic program, let us consider the example, Simplex #1, in which the variables A , M_0 , \underline{b} , and \underline{c} of the above model are assigned the same values as in Basis #1 (see Section 4.6). If we use a Monte Carlo simulation of 19000 such problems we find that the average value of z is 20.05181 and the variance of z is 2.08087 (see Table 5-2).

TABLE 5-2
Summary of Simplex #1 Simulation

Statistics	Number of Simplexes Solved				
	Monte Carlo			Sampling	
	1000	15000	19000	729	15625
mean z	20.14952	20.03670	20.05181	20.03091	20.02029
var. z	2.01308	2.05796	2.08087	2.17065	2.70947
mean x_1	0.10550	0.09515	0.09491	0.00000	0.12838
mean x_2	10.02202	9.97078	9.97845	10.01619	9.95965
mean x_3	4.68516	4.73671	4.73369	4.98345	4.64656
mean x_4	0.00010	0.00010	0.00008	0.00000	0.00002
var. x_1	0.21278	0.18976	0.18852	0.00000	0.24843
var. x_2	0.55056	0.55700	0.56189	0.53432	0.60255
cov(x_1, x_2)	-0.10045	-0.08992	-0.08877	0.00000	-0.11180
rel.freq.					
(x_1, x_2)	0.052	0.0488	0.0488	0.0000	0.1994
(x_2, x_3)	0.946	0.9509	0.9509	1.0000	0.7974
(x_2, x_4)	0.002	0.0003	0.0003	0.0000	0.0032

TABLE 5-3

Generated Distribution of z Using The Monte Carlo
Method And 19000 Simplexes

Upper Bound T	Rel.Freq. T	Cum.Rel. Freq. T	Upper Bound T	Rel.Freq. T	Cum.Rel. Freq. T
16.0000	0.0014	0.0014	20.0000	0.0345	0.4925
16.1250	0.0007	0.0021	20.1250	0.0336	0.5261
16.2500	0.0005	0.0026	20.2500	0.0323	0.5584
16.3750	0.0009	0.0036	20.3750	0.0346	0.5931
16.5000	0.0011	0.0047	20.5000	0.0319	0.6250
16.6250	0.0015	0.0062	20.6250	0.0332	0.6582
16.7500	0.0025	0.0087	20.7500	0.0314	0.6896
16.8750	0.0028	0.0115	20.8750	0.0324	0.7220
17.0000	0.0031	0.0145	21.0000	0.0249	0.7470
17.1250	0.0041	0.0186	21.1250	0.0257	0.7727
17.2500	0.0045	0.0231	21.2500	0.0258	0.7985
17.3750	0.0051	0.0282	21.3750	0.0236	0.8222
17.5000	0.0090	0.0372	21.5000	0.0220	0.8441
17.6250	0.0078	0.0449	21.6250	0.0192	0.8633
17.7500	0.0089	0.0539	21.7500	0.0164	0.8797
17.8750	0.0103	0.0642	21.8750	0.0148	0.8945
18.0000	0.0119	0.0761	22.0000	0.0142	0.9087
18.1250	0.0116	0.0876	22.1250	0.0124	0.9211
18.2500	0.0163	0.1040	22.2500	0.0119	0.9330
18.3750	0.0173	0.1212	22.3750	0.0103	0.9433
18.5000	0.0191	0.1404	22.5000	0.0092	0.9525
18.6250	0.0199	0.1603	22.6250	0.0064	0.9590
18.7500	0.0225	0.1828	22.7500	0.0066	0.9655
18.8750	0.0241	0.2069	22.8750	0.0061	0.9716
19.0000	0.0288	0.2357	23.0000	0.0052	0.9767
19.1250	0.0273	0.2631	23.1250	0.0046	0.9813
19.2500	0.0327	0.2958	23.2500	0.0034	0.9847
19.3750	0.0316	0.3274	23.3750	0.0033	0.9880
19.5000	0.0312	0.3586	23.5000	0.0022	0.9902
19.6250	0.0328	0.3914	23.6250	0.0019	0.9921
19.7500	0.0329	0.4243	23.7500	0.0014	0.9934
19.8750	0.0337	0.4580	+∞	0.0066	1.0000

TABLE 5-4

Generated Marginal Distribution of x_1 Using The Monte Carlo
Method And 19000 Simplexes

Upper Bound T	Rel.Freq. T	Cum.Rel. Freq. T	Upper Bound T	Rel.Freq. T	Cum.Rel. Freq. T
0.0000	0.9512	0.9512	1.0000	0.0005	0.9532
0.0625	0.0000	0.9512	1.0625	0.0004	0.9536
0.1250	0.0000	0.9512	1.1250	0.0005	0.9541
0.1875	0.0000	0.9512	1.1875	0.0008	0.9548
0.2500	0.0000	0.9512	1.2500	0.0009	0.9558
0.3125	0.0000	0.9512	1.3125	0.0012	0.9570
0.3750	0.0000	0.9512	1.3750	0.0014	0.9583
0.4375	0.0001	0.9513	1.4375	0.0012	0.9595
0.5000	0.0000	0.9513	1.5000	0.0011	0.9606
0.5625	0.0000	0.9513	1.5625	0.0017	0.9623
0.6250	0.0001	0.9514	1.6250	0.0016	0.9639
0.6875	0.0001	0.9515	1.6875	0.0018	0.9657
0.7500	0.0003	0.9518	1.7500	0.0024	0.9681
0.8125	0.0001	0.9518	1.8125	0.0019	0.9700
0.8750	0.0002	0.9521	1.8750	0.0018	0.9718
0.9375	0.0006	0.9526	$+\infty$	0.0282	1.0000

TABLE 5-5

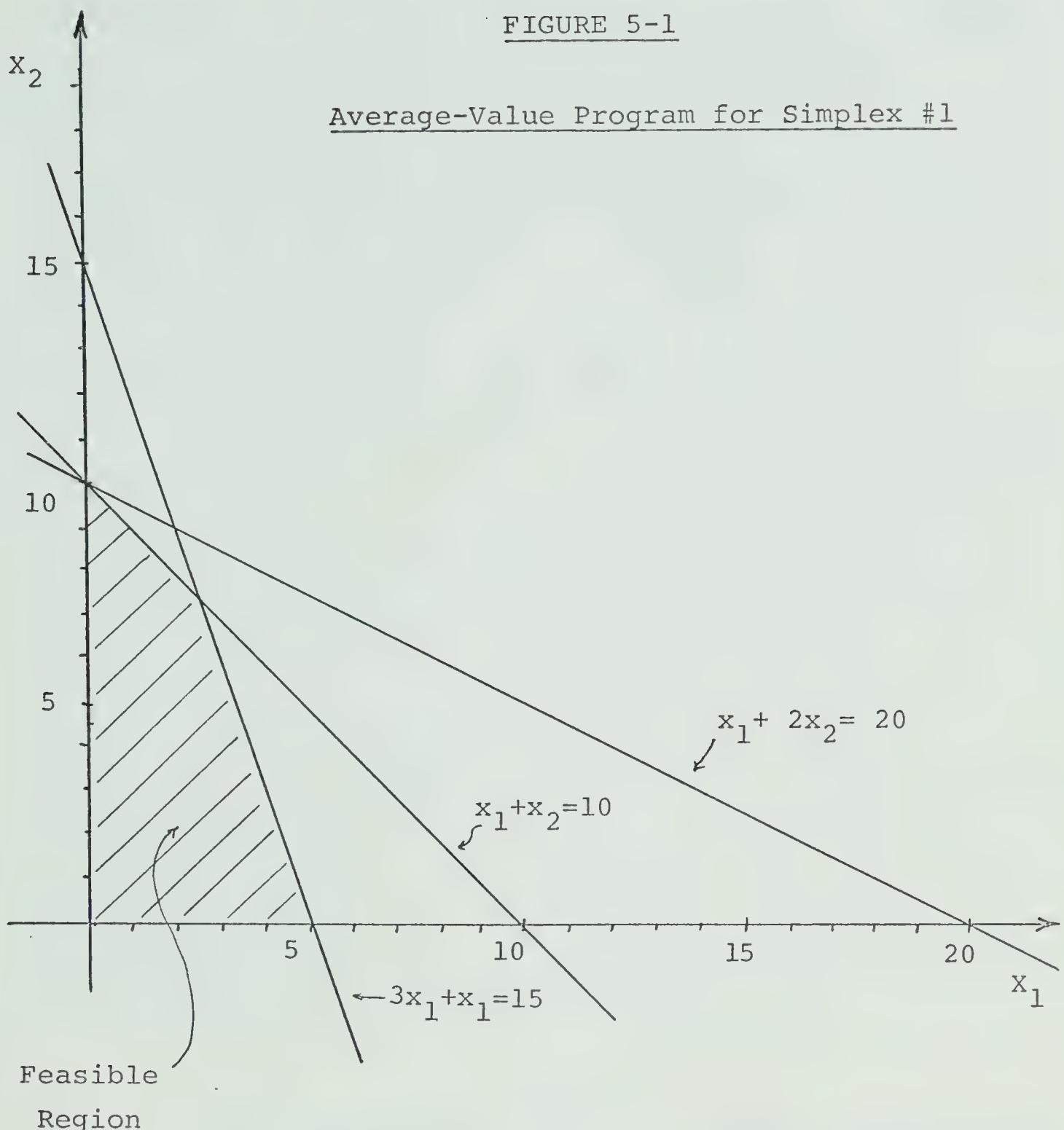
Generated Marginal Distribution of x_2 Using The Monte Carlo Method And 19000 Simplexes

Upper Bound	Rel.Freq. T	Cum.Rel. Freq. T	Upper Bound	Rel.Freq. T	Cum.Rel. Freq. T
8.0000	0.0054	0.0054	10.0000	0.0337	0.5128
8.0625	0.0009	0.0063	10.0625	0.0333	0.5461
8.1250	0.0012	0.0075	10.1250	0.0310	0.5771
8.1875	0.0017	0.0092	10.1875	0.0337	0.6107
8.2500	0.0017	0.0109	10.2500	0.0307	0.6415
8.3125	0.0026	0.0136	10.3125	0.0323	0.6737
8.3750	0.0031	0.0166	10.3750	0.0306	0.7044
8.4375	0.0038	0.0204	10.4375	0.0313	0.7356
8.5000	0.0038	0.0242	10.5000	0.0239	0.7596
8.5625	0.0048	0.0291	10.5625	0.0246	0.7842
8.6250	0.0058	0.0348	10.6250	0.0241	0.8083
8.6875	0.0058	0.0407	10.6875	0.0224	0.8306
8.7500	0.0102	0.0508	10.7500	0.0207	0.8514
8.8125	0.0087	0.0596	10.8125	0.0184	0.8697
8.8750	0.0098	0.0694	10.8750	0.0157	0.8855
8.9375	0.0116	0.0809	10.9375	0.0141	0.8996
9.0000	0.0124	0.0934	11.0000	0.0135	0.9131
9.0625	0.0130	0.1063	11.0625	0.0122	0.9252
9.1250	0.0173	0.1236	11.1250	0.0114	0.9366
9.1875	0.0181	0.1416	11.1875	0.0097	0.9463
9.2500	0.0199	0.1616	11.2500	0.0089	0.9553
9.3125	0.0209	0.1825	11.3125	0.0061	0.9613
9.3750	0.0228	0.2054	11.3750	0.0063	0.9676
9.4375	0.0245	0.2298	11.4375	0.0057	0.9733
9.5000	0.0288	0.2587	11.5000	0.0051	0.9783
9.5625	0.0271	0.2857	11.5625	0.0043	0.9826
9.6250	0.0325	0.3182	11.6250	0.0032	0.9858
9.6875	0.0316	0.3498	11.6875	0.0032	0.9890
9.7500	0.0311	0.3808	11.7500	0.0021	0.9911
9.8125	0.0322	0.4130	11.8125	0.0018	0.9929
9.8750	0.0324	0.4454	11.8750	0.0013	0.9942
9.9375	0.0336	0.4791	+∞	0.0058	1.0000

TABLE 5-6

Generated Marginal Distribution of x_3 Using the Monte Carlo Method And 19000 Simplexes

Upper Bound	Rel.Freq. T	Cum.Rel. Freq. T	Upper Bound	Rel.Freq. T	Cum.Rel. Freq. T
0.5000	0.0496	0.0496	4.5000	0.0344	0.3845
0.6250	0.0003	0.0499	4.6250	0.0314	0.4159
0.7500	0.0004	0.0503	4.7500	0.0333	0.4492
0.8750	0.0003	0.0505	4.8750	0.0320	0.4812
1.0000	0.0005	0.0510	5.0000	0.0362	0.5174
1.1250	0.0004	0.0514	5.1250	0.0354	0.5528
1.2500	0.0011	0.0525	5.2500	0.0362	0.5890
1.3750	0.0013	0.0538	5.3750	0.0351	0.6241
1.5000	0.0015	0.0553	5.5000	0.0352	0.6593
1.6250	0.0015	0.0568	5.6250	0.0343	0.6936
1.7500	0.0021	0.0589	5.7500	0.0333	0.7268
1.8750	0.0021	0.0610	5.8750	0.0289	0.7558
2.0000	0.0036	0.0646	6.0000	0.0312	0.7870
2.1250	0.0046	0.0693	6.1250	0.0264	0.8134
2.2500	0.0033	0.0725	6.2500	0.0245	0.8379
2.3750	0.0058	0.0783	6.3750	0.0219	0.8598
2.5000	0.0059	0.0842	6.5000	0.0232	0.8831
2.6250	0.0079	0.0922	6.6250	0.0173	0.9003
2.7500	0.0077	0.0998	6.7500	0.0171	0.9174
2.8750	0.0099	0.1098	6.8750	0.0150	0.9324
3.0000	0.0109	0.1207	7.0000	0.0125	0.9448
3.1250	0.0130	0.1337	7.1250	0.0106	0.9554
3.2500	0.0137	0.1474	7.2500	0.0080	0.9634
3.3750	0.0160	0.1634	7.3750	0.0069	0.9703
3.5000	0.0166	0.1800	7.5000	0.0061	0.9764
3.6250	0.0181	0.1981	7.6250	0.0052	0.9815
3.7500	0.0206	0.2186	7.7500	0.0042	0.9857
3.8750	0.0214	0.2401	7.8750	0.0032	0.9889
4.0000	0.0241	0.2642	8.0000	0.0031	0.9920
4.1250	0.0272	0.2914	8.1250	0.0019	0.9940
4.2500	0.0279	0.3193	8.2500	0.0017	0.9956
4.3750	0.0308	0.3502	+∞	0.0044	1.0000

FIGURE 5-1Average-Value Program for Simplex #1

Feasible
Region

$$\max_{x_i} (1 \quad 2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (1 \quad 2) \begin{pmatrix} 0 \\ 10 \end{pmatrix} = 20$$

subject to

$$\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 15 \\ 10 \end{pmatrix}$$

and

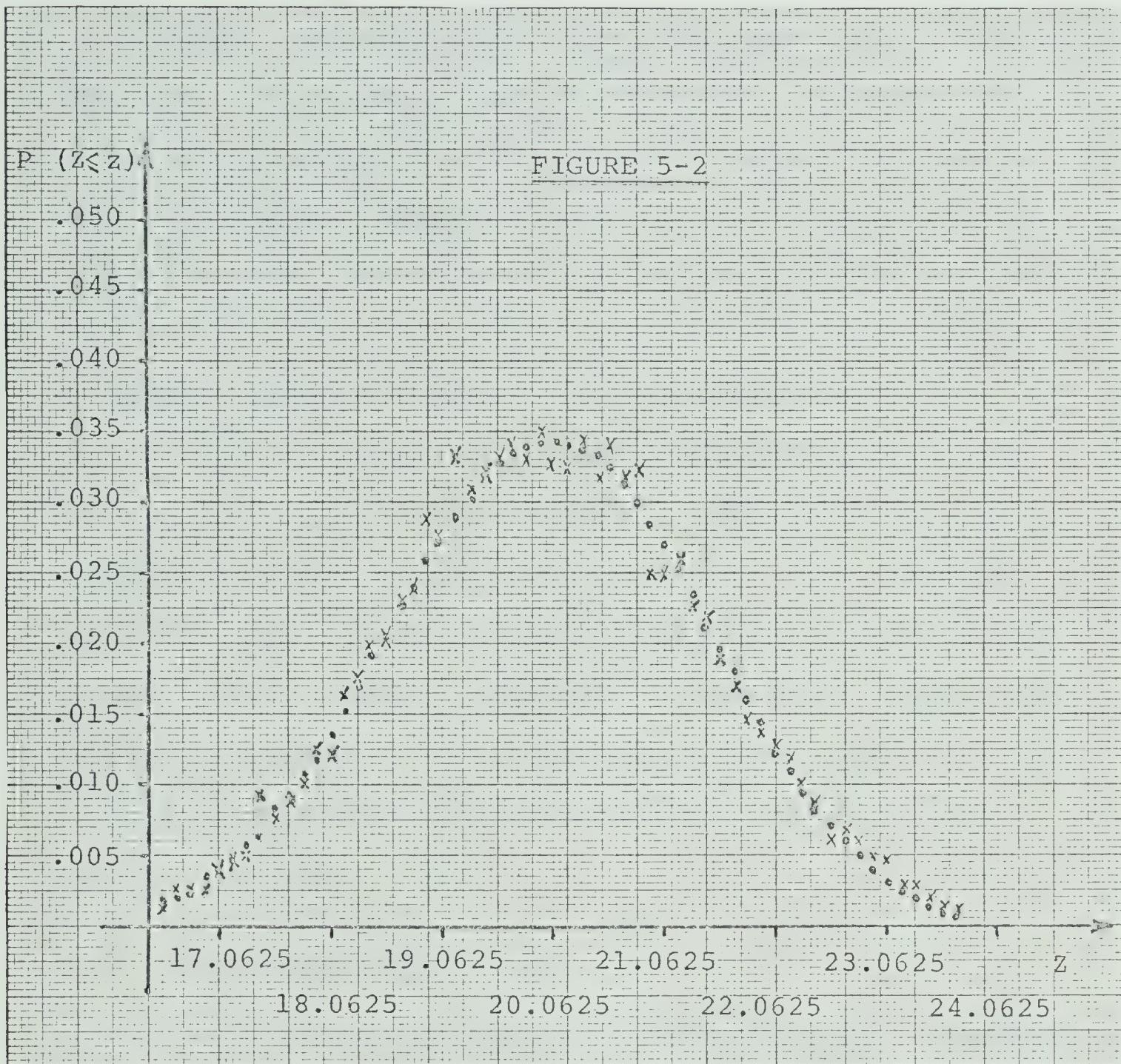
$$x_1 \geq 0, \quad x_2 \geq 0.$$

The average-value model of Simplex #1 is shown in Figure 5-1. The difference between the average optimal value of the stochastic objective function and the optimal value of the average-value objective function is small. However, the probability that $z \leq 19$ or $z \geq 21$ is approximately equal to 0.5, i.e., the optimal value of the stochastic model is not very stable. In other words, the probability that the absolute error, $|20-z|$, is greater than 1 is at least 0.5. The upper bound T in the frequency table of z , Table 5-3, refers to the upper bound of the frequency class, i.e., the first three frequency classes are $(-\infty, 16]$, $(16, 16.125]$, and $(16.125, 16.25]$. The second column contains the relative frequencies of the intervals. The third column contains the cumulative relative frequencies of the point T , e.g., the cumulative relative frequency of 17.75 is 0.0539. The corresponding information for x_1 , x_2 , x_3 , and x_4 is contained in Tables 5-4, 5-5, 5-6, and 5-7, respectively.

TABLE 5-7

Generated Distribution of x Using The Monte Carlo Method and 19000 Simplexes

Interval $(\alpha, \beta]$		Rel.Freq. $(\alpha, \beta]$	Cum. Rel. Freq. $(\alpha, \beta]$
α	β		
$-\infty$	0.0000	0.9997	0.9997
0.0000	0.0625	0.0001	0.9998
0.0625	0.1875	0.0001	0.9999
0.1875	0.3125	0.0001	0.9999
0.3125	1.0625	0.0001	1.0000



Comparison of the Normal Density and the
Generated Density for z

\times Simplex #1 $E[z] = 20.03670$, $\text{Var}(z) = 2.05796$
using 15000 points

. Normal Variate, $\mu = 20.03670$, $\sigma^2 = 2.05796$

While we cannot prove that the resulting distribution of z is normal, it is interesting to compare the observed distribution with a normal variate whose mean and variance are equal to those of the observed statistic. This was done for Simplex #1 using 15000 random simplexes and the comparison is shown in Figure 5-2. It shows that a normal approximation for z is reasonable.

Since Tintner always uses the Sampling technique to simulate his stochastic models, Simplex #1 was also simulated using the Sampling technique. The first two approximations given in Table 4-7 were used and the comparison of the Monte Carlo and the Sampling results are shown in Table 5-2. The comments made earlier about the Sampling technique, regarding the bias and size, are again applicable.

5.4 Preference Functionals

We mentioned in Chapter 4 that stochastic programming could be used in problems involving decision-making under risk, i.e., there may exist several possible probability density functions of A , b , and c

$$\{\Pr(A, b, c | \underline{\theta}) | \underline{\theta} \in \Omega\} , \quad (5.4-1)$$

and we might want to choose the density which yields the maximum expected value of z ,

$$\max_{\underline{\theta} \in \Omega} E[z | \underline{\theta}] . \quad (5.4-2)$$

For example, if we are given two possible density functions for Simplex #1, we would want to use the model which has the maximum expected value of the objective function. Let the density defined by Simplex #1 be

$$\Pr(A, \underline{b}, \underline{c} | \underline{\theta}_1) . \quad (5.4-3)$$

For an alternative model, Simplex #2, we define a model which is the same as Simplex #1 except that

$$\text{Var}(A) = \begin{bmatrix} 0.01 & 0.005 \\ 0.05 & 0.001 \end{bmatrix} , \quad (5.4-4)$$

$$\text{Var}(\underline{b}) = (0.64, 0.16) ; \quad (5.4-5)$$

and denote its density by

$$\Pr(A, \underline{b}, \underline{c} | \underline{\theta}_2) . \quad (5.4-6)$$

Both models were simulated and the results of the two experiments are shown in Table 5-8.

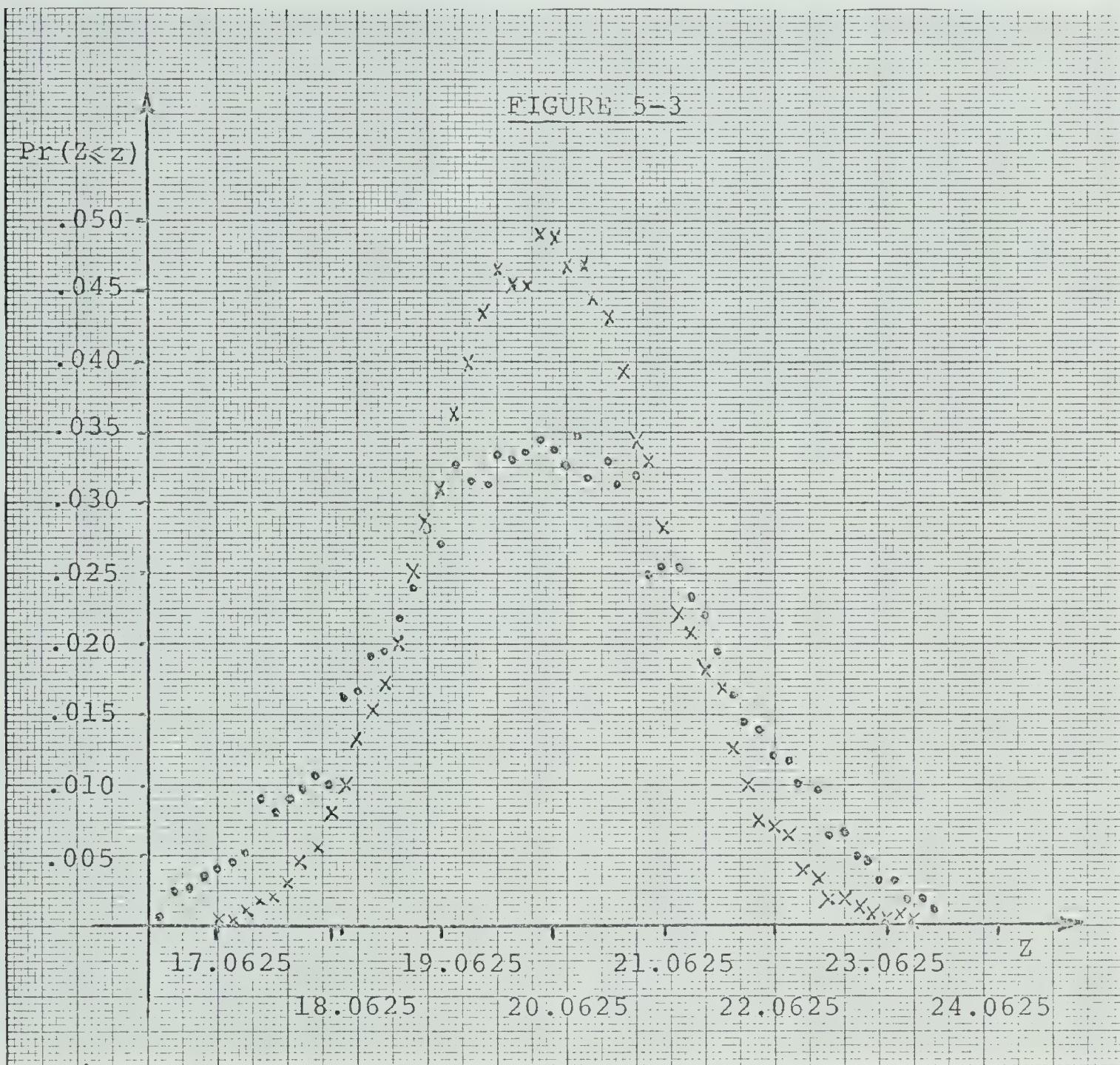
TABLE 5-8
An Alternative To Simplex #1 Generated
From 19000 Random Simplexes

Statistics	Simplex #1	Simplex #2
mean z	20.05181	20.02208
variance z	2.08087	1.03841
mean x_1	0.09491	0.02531
mean x_2	9.97845	9.99839
mean x_3	4.73369	4.91831
mean x_4	0.00008	0.0002
variance x_1	0.18852	0.05151
variance x_2	0.56189	0.27255
cov(x_1, x_2)	-0.08877	-0.02580
rel.freq.(x_1, x_2)	0.0488	0.0129
rel.freq.(x_2, x_3)	0.9509	0.9870
rel.freq.(x_2, x_4)	0.0003	0.0001

Thus, if we are to select the model which has the greatest expected value of z , then we would select Simplex #1 since $20.05181 > 20.02208$, i.e.,

$$\max_{\underline{\theta} \in \Omega} E[z | \underline{\theta}] = E[z | \underline{\theta}_1], \quad \Omega = \{\underline{\theta}_1, \underline{\theta}_2\}. \quad (5.4-7)$$

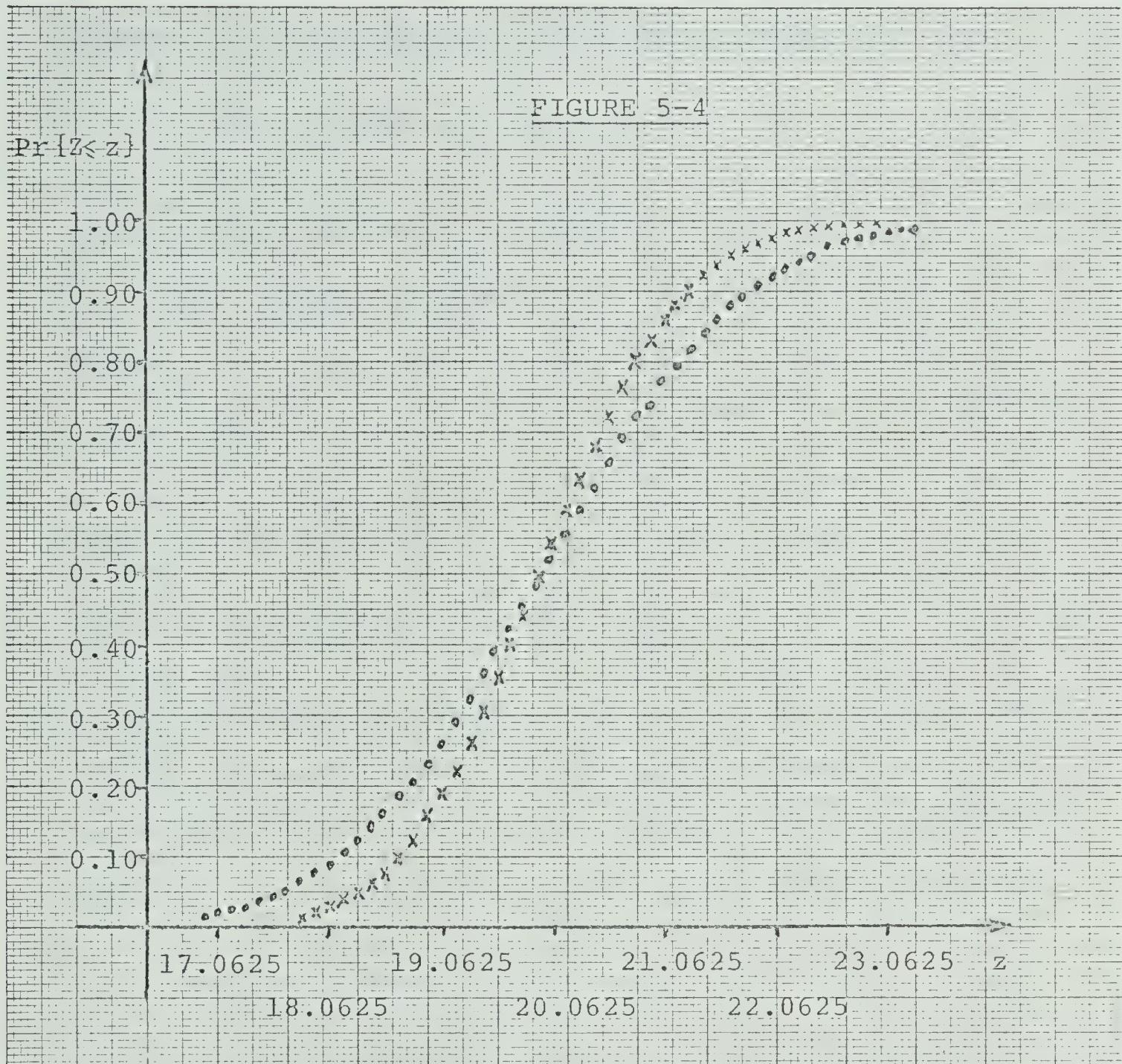
FIGURE 5-3



Generated Density Functions for z

in Simplex #1 and Simplex #2

- Simplex #1 $E[z] = 20.05181$, $\text{Var}(z) = 2.08087$
- ✗ Simplex #2 $E[z] = 20.02208$, $\text{Var}(z) = 1.03841$



Generated Distribution Functions for z

in Simplex #1 and Simplex #2

\times Simplex #1 $E[z] = 20.05181, \quad \text{Var}(z) = 2.08087$

\circ Simplex #2 $E[z] = 20.02208, \quad \text{Var}(z) = 1.03841$

However, if the preference functional is of the form

$$\min_{\underline{\theta}} \text{Var}(z | \underline{\theta}) , \quad \underline{\theta} \in \Omega , \quad (5.4-8)$$

then we would select the alternative model.

The preference functional may even have the form

$$\max_{\underline{\theta}_i} \alpha_i , \quad i=1,2 , \quad (5.4-9)$$

where α_i is such that

$$\Pr(z \leq \alpha_i | \underline{\theta}_i) = 0.05 , \quad (5.4-10)$$

i.e., we choose the distribution with the greatest value of the lower 5% point. The relative frequencies of the two distributions can be compared by viewing Figures 5-3 and 5-4. We find that we would select Simplex #2 since

$$\max_{\underline{\theta}_i} \alpha_i = \max\{\alpha_1, \alpha_2\} = \alpha_2$$

$$\max_{\underline{\theta}_i} \alpha_i \approx \max \{17.75, 18.375\} = 18.375 .$$

The above illustrations have indicated how stochastic programming can be used in decision-making problems under risk.

5.5 Active and Passive Stochastic Programming

A variation of the stochastic linear programming model introduced by Tintner (1960) will be discussed in this section. Tintner named the model defined by (5.2-1), (5.2-2), (5.2-3), and (5.2-4) the passive stochastic linear programming model. It is passive in the sense that the decision variables are chosen after the values of the random variables are observed. Thus the passive model resembles the wait-and-see model of Chapter 2. The alternative model, referred to as the active stochastic linear programming model, is defined as

$$\max_{\underline{x}} \underline{c}' \underline{x} \quad (5.5-1)$$

subject to

$$A \mathcal{D}(\underline{x}) \leq \mathcal{D}(\underline{b}) U , \quad (5.5-2)$$

and

$$\underline{x} \geq \underline{0} , \quad (5.5-3)$$

where A , \underline{b} , and \underline{c} are the same as in the passive model. The operator \mathcal{D} is the diagonalization operator defined previously. The matrix $U = [u_{ij}]$ is an m by n matrix such that

$$0 \leq u_{ij} \leq 1, \quad i=1,2,\dots,m, \quad j=1,2,\dots,n, \quad (5.5-4)$$

and

$$\sum_{j=1}^n u_{ij} = 1, \quad i=1,2,\dots,m. \quad (5.5-5)$$

The quantity u_{ij} represents the proportion of the i^{th} resource, b_i , allocated to the j^{th} activity, $A_{;j}$. The allocation matrix, U , is assigned values in accordance with (5.5-4) and (5.5-5) before the random events occur. Thus the active model resembles the here-and-now model of Chapter 2. In the active model, the distribution of z is a function of U as well as of θ . Hence, if a preference functional is to be maximized for the active model, it can be maximized over the set of possible allocation matrices, $\{U\}$. The active formulation does not permit a simpler analytic solution. To solve an active model, we must again simulate the model and use the results to predict the "real world" phenomena. The paper by Tintner and Raghavan (1965) presents a fairly good example of such a simulation.

Some theorems on the relationship between the active and passive stochastic linear programs are presented by Senguta, Tintner, and Morrison (1963) and by Senguta, Tintner, and Millham (1963). These theorems are similar to those presented in Section 2.5 and as in Section 2.5

they find that the optimal value of the active model can never be better than the optimal value of the passive model (see (2.5-8)). In another paper, Tintner, Millham, and Senguta (1962) extend the duality theorem of deterministic linear programming to stochastic linear programming.

Before concluding this study, we should again mention that although the active-passive classification provides an alternative formulation for the stochastic linear programming problem, it does not simplify the solution to this problem.

CHAPTER 6CONCLUSION

If a "real world" problem can be modelled as a linear programming model under uncertainty or a chance-constrained model, then we have shown (Chapters 2 and 3) that there are acceptable techniques for solving these models. If on the other hand, a stochastic linear programming model must be used, then the easiest approximation to the solution is found by simulating the model. Although it may be theoretically possible to obtain an analytic solution or an algebraic approximation for the decision vector in the stochastic linear programming model, it appears that the best alternative to analyse the problem is to simulate the stochastic model. However, if the problem that we wish to solve can be expressed as a system of stochastic linear equations which satisfy the conditions of Section 4.7, then we can either use the approximation proposed in Chapter 4 or simulate the system of stochastic linear equations.

The first order estimator \hat{x} proposed in this thesis is equivalent to the estimator \hat{x}_t introduced by Tintner (1955) except when both A and b are random variables. The development of \hat{x} is more rigorous and is easier to

extend to higher order approximations than is the case for \hat{x}_t . Also using \hat{x} as an estimator, we were able to derive in Section 4.7 a sufficient condition for x in any particular stochastic model. We also verified that \hat{x} is a good approximation for x by simulating a number of different systems of stochastic linear equations.

Furthermore, this study has shown that APL in its present implementation is not suitable for large or lengthy simulations. However, by using the APL random number generator in assembler form, we have verified that this random number generator is satisfactory for this type of simulation, i.e., the simulations confirm the results of Appendix 1.

While it is hoped that this study has drawn together many existing techniques and added a few of its own, there are still several unanswered questions. One of these questions is how useful are stochastic linear programs in "real world" situations. Although the average value of the optimal solution of a stochastic linear programming model may be very close to the optimal solution of the corresponding average-value model, the optimal solution of the stochastic model may be very unstable. Hence researchers should be encouraged to incorporate stochastic variations into their models so that the merits of the stochastic linear programs may be assessed.

Probably the most interesting unanswered question is the determination of the probability density function, f_n , of the n^{th} order estimator,

$$\hat{x}_n = \{ I + \sum_{j=1}^n (-1)^j (M A)^j \} M b . \quad (6-1)$$

Once f_n has been determined then it may be possible to show that $f_n \rightarrow f$ as $n \rightarrow \infty$ where f is the probability density function of x . Also, if the probability of any given basis being optimal were known, then it would be practical to find approximations of the type mentioned in Chapter 5. Furthermore, we could make the simulation of stochastic linear programs more efficient by using parametric programming techniques instead of always using the simplex algorithm to solve each linear program simulated.

As the previous pages have shown, the solution of the stochastic linear program is a complicated procedure, and until simpler procedures are found, the stochastic linear program will not become an important "real world" model. However, only the use of the stochastic linear program will determine its merit as a "real world" model and encourage further research.

REFERENCES

- Adams, W.S. (1967). Introduction to APL\360 Concepts.
Report, Department of Computing Science, University
of Alberta, Edmonton, Alberta.
- Babbar, M.M. (1955). "Distribution of Solutions of a Set
of Linear Equations with Applications to Linear
Programming". Journal of the American Statistical
Association, Vol. 50, pp. 854-869.
- Box, G.E.P., and M.E. Muller (1958). "A Note on the
Generation of Random Normal Deviates". Annals
of Mathematical Statistics, Vol. 28, pp. 610-611.
- Charnes, A., and W.W. Cooper (1963 a). "Deterministic
Equivalents for Optimizing and Satisficing under
Chance Constraints". Operations Research, Vol. 11,
pp. 18-39.
- Charnes, A., and W.W. Cooper (1962). "Chance Constraints
and Normal Deviates". Journal of the American
Statistical Association, Vol. 57, pp. 134-148.
- Charnes, A., and W.W. Cooper (1959). "Chance-Constrained
Programming". Management Science, Vol. 6, pp. 73-
80.
- Charnes, A., W.W. Cooper, and G.H. Symonds (1958).
"Cost Horizons and Certainty Equivalents An
Approach to Stochastic Programming of Heating
Oil". Management Science, Vol. 4, pp. 235-263.

- Charnes, A., W.W. Cooper, and G.L. Thompson (1964). "Critical Path Analysis Via Chance Constrained and Stochastic Programming". Operations Research, Vol. 12, pp. 460-470.
- Cochran, W.G. (1954). "Some Methods for Strengthening the Common Chi-Square Tests". Biometrics, Vol. 10, pp. 417-451.
- Cochran, W.G. (1952). "The χ^2 Test of Goodness of Fit". Annals of Mathematical Statistics, Vol. 23, pp. 315-345.
- Cramér, H. (1946). Mathematical Methods of Statistics. Princeton University Press, Princeton.
- Dantzig, G.B. (1955). "Linear Programming Under Uncertainty". Management Science, Vol. 1, pp. 197-206.
- Dantzig, G.B., and A. Madansky (1960). "On the Solution of Two-Stage Linear Programs Under Uncertainty". Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, pp. 165-176.
- Dantzig, G.B., and P. Wolfe (1960). "Decomposition Principle for Linear Programs". Operations Research, Vol. 8, pp. 101-111.
- Falkoff, A.D., and K.E. Iverson (1968). APL\ 360 : User's Manual. IBM Thomas J. Watson Research Center.

- Ferguson, A.R., and G.B. Dantzig (1956). "The Allocation of Aircraft to Routes - an example of linear programming under uncertain demand". Management Science, Vol. 3, pp. 45-73.
- Gorenstein, S. (1967). "Testing a Random Number Generator". Communications of A.C.M., Vol. 10, pp. 111-118.
- Hadley, G. (1963). Linear Programming. Addison-Wesley, Reading, Massachusetts.
- Hillier, F.S., and E.J. Liebermann (1967). Introduction to Operations Research. Holden-Day, San Francisco.
- Hogg, R.V., and A.T. Craig (1965). Introduction to Mathematical Statistics (second edition). MacMillan, New York.
- Hull, T.E., and A.R. Dobell (1962). "Random Number Generators". SIAM Review, No. 4, pp. 230-254.
- Hutchinson, D.W. (1966). "A New Uniform Pseudorandom Number Generator". Communications of the A.C.M., Vol. 9, pp. 432-433.
- Jansson, B. (1966). Random Number Generators. Victor Pettersons Bokindustri Aktiebolag.
- Keeping, E.S. (1962). Introduction to Statistical Inference. Van Nortrand, Princeton N.J.
- Kendall, M.G., and A. Stuart (1958). The Advanced Theory of Statistics, Vol. 1, pp. 270-271. Charles Griffin and Co., London.

Lewis, P.A., A.S. Goodman, and J.M. Miller (1969).

"A Pseudo-Random Number Generator for the System/360". IBM System Journal, Vol. 8, pp. 136-146.

Madansky, A. (1963). "Linear Programming Under Uncertainty". Recent Advances in Mathematical Programming. R.L. Graves and P. Wolfe editors. McGraw-Hill, New York.

Madansky, A. (1960). "Inequalities for Stochastic Linear Programming Problems". Management Science, Vol. 6, pp. 197-204.

Mangasarian, O.L., and J.B. Rosen (1964). "Inequalities for Stochastic Nonlinear Programming Problems". Operations Research, Vol. 12, pp. 143-154.

Miller, B.L., and H.M. Wagner (1965). "Chance Constrained Programming with Joint Constraints". Operations Research, Vol. 13, pp. 930-945.

Näslund, B. (1965). "Mathematical Programming Under Risk". The Swedish Journal of Economics, Vol. 18, pp. 241-253.

Sengupta, J.K., G. Tintner, and C. Millham (1963). "On Some Theorems of Stochastic Linear Programming with Applications". Management Science, Vol. 10, pp. 143-159.

- Sengupta, J.K., G. Tintner, and B. Morrison (1963). "Stochastic Linear Programming with Applications to Economic Models". Economica, Vol. 30, pp. 262-276.
- Symonds, G.H. (1968). "Chance-Constrained Equivalents of Some Stochastic Programming Problems". Operations Research, Vol. 16, pp. 1152-1160.
- Symonds, G.H. (1967). "Deterministic Solution for a Class of Chance-Constrained Programming Problems". Operations Research, Vol. 15, pp. 495-512.
- Thompson, G.L., W.W. Cooper, and A. Charnes (1963 b). "Characterizations by Chance-Constrained Programming". Recent Advances in Mathematical Programming. R.L. Graves and P. Wolfe editors. McGraw-Hill, New York.
- Tintner, G. (1960). "A Note on Stochastic Linear Programming". Econometrica, Vol. 28, pp. 490-495.
- Tintner, G. (1955). "Stochastic Linear Programming with Applications to Agricultural Economics". Second Symposium on Linear Programming, Vol. 1, pp. 197-228. H.A. Antosiewics editor. National Bureau of Standards, Washington, D.C.
- Tintner, G., C. Millham, and J.K. Sengupta (1962). "A Weak Duality Theory for Stochastic Linear Programming". Unternehmensforschung, pp. 1-8.

- Tintner, G., and N.S. Raghavan (1965). "Stochastic Linear Programming Applied to a Dynamic Planning Model for India". Monograph.
- Tintner, G., J.K. Sengupta, and V.Y. Rao (1962). "An Application of Stochastic Linear Programming to Development Planning". Metroeconomica, Vol. 14, pp. 25-41.
- Wagner, H.M. (1969). Principles of Operations Research with Applications to Managerial Decisions. Prentice-Hall, Englewood Cliffs, New Jersey.
- Wagner, H.M. (1958). "On the Distribution of Solutions in Linear Programming Problem". American Statistical Association Journal, Vol. 53, pp. 161-163.
- Wilson, R. (1966). "On Programming Under Uncertainty". Operations Research, Vol. 14, pp. 652-657.

APPENDIX 1

TESTING OF THE APL RANDOM NUMBER GENERATOR

A1.1 Introduction

The object of this appendix is to describe the testing of the APL random number generator in order to justify the use of APL for simulations. The only other testing of the APL random number generator appears in the paper by Lewis, Goodman, and Miller (1969) which uses Good's period, Fourier-transform, and run tests on the numbers generated by the APL random number generator. Although the tests used in this appendix are different, the conclusion is the same that the random number generator is very good. The paper by Gorenstein (1967) presents results for the GPSS (General Purpose Simulation System) random number generator, while the bibliography of this paper and the bibliography of Hull and Dobell (1962) and of Jansson (1966) contain references to other random number generators.

A random number generator is a device, which produces numbers in some nondeterministic way. The sequence of numbers appearing when a roulette wheel is spun a number of times is random since we do not know a priori which of the numbers will appear on any particular spin. This method

of generation could be simulated by the computer if we equipped it with a mechanical roulette wheel or stored a table of roulette wheel spins. These two methods are not used because the former makes it impossible to repeat an experiment, while the latter requires too much time and storage.

Instead, a recursive mathematical method is used to generate a sequence of numbers with certain properties. Since any number in the sequence may be determined provided its predecessor and the method are known, the numbers are called "pseudo-random numbers". The particular recursive method used in APL is due to Lehmer (see Hutchinson (1966)) and may be defined as follows.

For parameters k and x_1 we may define a sequence of numbers x_1, x_2, \dots , as

$$x_{n+1} = k x_n \pmod{p}, \quad n=1, 2, \dots \quad (\text{Al.1-1})$$

or in APL notation

$$X \leftarrow P | K \times X .$$

Such a generator is referred to as a multiplicative congruential pseudo-random number generator. The parameter k is the multiplier, x_1 is the seed, and p is the modulus. The characteristics of the resulting string of numbers is

dependent upon the particular set of parameters chosen as is shown by the following example from Hillier and Lieberman (1967, p.446).

If $x_0 = 3$, $k = 7$, $p = 10$, and

$$x_{n+1} = 7 x_n \pmod{p}, \quad n=1,2,\dots,$$

then the resulting string of numbers is

$$1, 7, 9, 3, 1, 7, 9, 3, 1 \dots .$$

Notice this particular combination of parameters yields a sequence which repeats itself every four digits. It would seem reasonable, that among other properties of the generator, we would want the cycle length as long as possible. The cycle length is determined by the values of k and p according to some number theoretic theorems. However, these theorems do not yield a unique set of values for k and p , but rather they suggest optimum properties of k and p . First p should be prime and as large as possible. Since the word length of the APL interpreter for the IBM System / 360 is thirty-two bits, the maximum value for p is $2^{31} - 1$, which conveniently happens to be prime. Then k should be a positive primitive root of p and should be close to \sqrt{p} . Lewis, Goodman, and Miller (1969) show that $7^5 = 16807$ is such a value for k . Thus

the APL random number generator is

$$x_{n+1} = x_n \cdot 7^5 \pmod{2^{31} - 1}, \quad n=1, 2, \dots, \quad (\text{A1.1-2})$$

or

$$X \leftarrow (-1 + 2 * 31) | X \times 7 * 5 .$$

This particular generator has cycle length of $2^{31} - 2$.

The seed in a clear workspace (the workspace obtained by the)CLEAR command) is the same as the multiplier, 16807. APL then saves the last random number generated and uses this number as x_n (or X) for the next random number generation. A)SAVE command will store the last random number generated for the seed when the workspace is reloaded ()LOAD).

We now have obtained a random generator (A1.1-2) with a maximum cycle length. What other properties do we want of a generator? First, we want a simple, short, and fast generating procedure, which obviously (A1.1-2) is. The paper by Lewis, Goodman, and Miller (1969) mentions an upper bound of approximately 31.2 microseconds per random number call in FORTRAN for a System / 360, Model 67. Secondly, the particular method of generation must be reproducible, a necessary condition if we are to be able to repeat simulation experiments. We can do this

with the method (A1.1-2). Lastly, the numbers generated must be statistically acceptable, i.e., they must appear to be random numbers. The particular statistical properties which are most important will depend upon the particular application employing the random numbers. Usually the random numbers are used to simulate a given distribution. The question then is how well do these random numbers simulate the distribution in question.

A1.2 The APL Random Number Generator

Before considering the specific tests, let us first look at the testing procedure used in APL. The code used in the interpreter to implement the generator just described is not accessible to the APL user. The user has access only to the operator ? which he may use either as a monadic or dyadic operator. We are interested only in the monadic ?. If K is a positive integer scalar, then ?K produces one number drawn at random from the set $\{1, 2, \dots, K\}$ (, i.e., $1K$). Thus the numbers generated by ?K should be values from the discrete uniform distribution

$$p(x) = 1/x , \quad \text{for } x=1, 2, \dots, K ,$$

$$= 0 , \quad \text{otherwise} .$$

It is the acceptability of the random number generator in this form in which we are interested, since any simulation in APL must ultimately be defined in terms of the monadic ? . The value of K used was 10, and the number $?K$ was added to $\neg 1$ in order to produce a value in the set $\{0,1,2,\dots,9\}$, i.e., the particular form of the generator was $\neg 1 + ? 10$. The starting value chosen as the seed was 16807. In other words, the tests were performed on the random numbers that would be found by working in a clear workspace. The APL function written to test the random numbers was *TEST* with the header *N TEST M* where N and M are positive integer scalars. The listing of this function may be found at the end of this appendix. This function is self-contained except for the global variable *PRINT* (a positive integer scalar) and the subroutine *RND* (a rounding routine, see Adams (1967), pp.6-7). The function tests $N \times M$ random numbers as well as each subset of $PRINT \times M$ random numbers where M is less than or equal to 1000 (because of limited space) and a multiple of 4 (because of the Poker Test). The function types a report on each subset of $PRINT \times M$ random numbers as well as a summary on all $N \times M$ random numbers. The particular values of (N, M, PRINT) used were (10, 1000, 1), (50, 1000, 5), and (100, 1000, 10). Only the last results of each run will be reported here, i.e., the results of the first 10000, 50000, and 100000

random numbers. Table A1-1 gives some timing estimates for the function *TEST*.

TABLE A1-1

Timing Estimates For The Function 'TEST'

Time Classes	N × M = 50000 Random Numbers	N × M = 100000 Random Numbers
Average CPU time per report typed	6 min. 22 sec.	13 min. 18 sec.
Total CPU time for TEST	63 min. 40 sec.	133 min.
Average terminal response time per report typed	28 min. 55 sec.	39 min. 1 sec.
Total terminal response time for TEST	4 hr. 49 min. 11 sec.	5 hr. 12 min. 4 sec.
Average CPU time for generation per number	0.00013 sec.	0.00013 sec.
Total CPU for random number generation	6.667 sec.	12.95 sec.

A1.3 Description of Statistical Tests

In this section we shall discuss the tests that were used to test the monadic operator ? . The six tests were the frequency test, odd and even test, poker test, gap test, coupon test, and serial test. Each will be described briefly

a) Frequency Test

The assumption that the numbers produced by $\neg 1 + ? 10$ are uniformly distributed can be tested statistically by using a frequency test. The number of occurrences of each of the digits $0, 1, \dots, 9$ is counted in the string D of M random numbers produced by TEST where

$$D = (d_i) \quad \text{and} \quad d_i \in \{0, 1, \dots, 9\}, \quad i=1, 2, \dots, M \quad (\text{A1.3-1})$$

or

$$D \leftarrow \neg 1 + ? M \rho 10$$

The frequency of each digit can then be used to construct a frequency vector F such that

$$F = (f_j), \quad i=0, 1, \dots, 9 \quad (\text{A1.3-2})$$

or in APL

$$F \leftarrow +/(0, 19) \circ . = D$$

where f_j is the number of j's in the vector D. This result is contained in the first frequency table typed by TEST which consists of a 10 by 4 table in which the first column contains the digit, the second column the frequency F, the third column the relative frequency ($f_0/M, \dots, f_9/M$) (or $F \div M$), and the fourth column the cumulative relative frequency. The results of this test for the three runs are shown in Tables Al-2, Al-3, and Al-4.

TABLE Al-2
Frequency Table Produced By 'TEST'
For 10000 Random Numbers

Digit	Frequency	Relative Frequency	Cumulative Relative Frequency
0	992	0.0992	0.0992
1	1007	0.1007	0.1999
2	998	0.0998	0.2997
3	958	0.0958	0.3955
4	1001	0.1001	0.4956
5	1049	0.1049	0.6005
6	989	0.0989	0.6994
7	964	0.0964	0.7958
8	1026	0.1026	0.8984
9	1016	0.1016	1.0000

TABLE A1-3Frequency Table Produced By 'TEST'For 50000 Random Numbers

Digit	Frequency	Relative Frequency	Cumulative Relative Frequency
0	4975	0.0995	0.0995
1	4981	0.09962	0.19912
2	4925	0.0985	0.29762
3	5019	0.10038	0.398
4	5029	0.10058	0.49858
5	5163	0.10326	0.60184
6	5032	0.10064	0.70248
7	4930	0.0986	0.80108
8	4979	0.09958	0.90066
9	4967	0.09934	1.00000

TABLE A1-4Frequency Table Produced By 'TEST'For 100000 Random Numbers

Digits	Frequency	Relative Frequency	Cumulative Relative Frequency
0	10046	0.10046	0.10046
1	10016	0.10016	0.20062
2	9863	0.09863	0.29925
3	9878	0.09878	0.39803
4	10012	0.10012	0.49815
5	10286	0.10286	0.60101
6	9931	0.09931	0.70032
7	9955	0.09955	0.79987
8	10118	0.10118	0.90105
9	9895	0.09895	1.00000

Theoretically, the numbers generated should have the probability density function

$$\begin{aligned} \Pr(X=x) &= 0.1 , \quad x=0,1,\dots, 9 , \\ &= 0 , \quad \text{otherwise.} \end{aligned} \quad (\text{A1.3-3})$$

If the numbers are random then they should have the above probability density function and thus it is possible to check the moments of this distribution against the appropriate sample statistics. The first moment or mean is

$$\mu = E[X] = \sum_x x p(x) = 4.5 \quad (\text{A1.3-4})$$

or

$$(+/X) \div 10 \equiv 4.5 \quad \text{where } X \leftarrow 0, 19$$

which corresponds to the sample mean calculated from D as follows

$$m = \frac{1}{M} \sum_{i=1}^M d_i \quad \text{or} \quad \frac{1}{M} \sum_{j=0}^9 j f_j \quad (\text{A1.3-5})$$

or written in APL

$$(+/D) \div \rho D \quad \text{or} \quad F +.\times 0, 19 .$$

It was necessary because of space limitations to calculate the sample statistics from the frequency vector. No grouping error is incurred by doing this since each frequency

class consists of only one point. A statistical test of the sample mean can be developed since the sampling distribution of m is approximately normal because the sample size is large and the population variance is finite. Thus m can be considered as a value of a normal variate with mean $\mu = 4.5$ and standard deviation $\sigma/\sqrt{N \times M} = 2.87228/\sqrt{N \times M}$. One such test for the sample mean could be: if m is in the interval

$$\left(\mu - \frac{a\sigma}{\sqrt{N \times M}}, \mu + \frac{a\sigma}{\sqrt{N \times M}} \right)$$

where a is determined by the $\alpha\%$ confidence level such that

$$\Pr \left(-a \leq \frac{M - \mu}{\sigma/\sqrt{N \times M}} \leq a \right) = \frac{\alpha}{100}, \quad (\text{Al.3-6})$$

and would accept the hypothesis that the mean of the population is 4.5 at the $\alpha\%$ confidence level. Since in this appendix we will arbitrarily work with $\alpha=90$, a must be 1.645. Using these values for α and a in (Al.3-6), we can determine the following confidence intervals in Table Al-5.

TABLE A1-5

Confidence Intervals For The Sample
Mean In The Frequency Test

Confidence Level	Number of Random Numbers	Confidence Intervals	Sample Mean
90%	10000	(4.453, 4.547)	4.516
90%	50000	(4.479, 4.521)	4.501
90%	100000	(4.485, 4.515)	4.501

The value of the second moment about zero is

$$\mu'_2 = E[X^2] = \sum_x x^2 p(x) = 28.5 \quad (\text{A1.3-7})$$

or calculated in APL

$$+/ (X * 2) \div 10 \equiv 28.5$$

while the corresponding sample statistic is

$$m'_2 = \frac{1}{M} \sum_{i=1}^M d_i^2 \quad \text{or} \quad \frac{1}{M} \sum_{j=0}^9 j^2 f_j \quad (\text{A1.3-8})$$

or

$$F + . \times (0, 19) * 2 \quad .$$

The third moment about zero is

$$\mu'_3 = E[x^3] = \sum_x x^3 p(x) = 202.5 \quad (\text{Al.3-9})$$

or

$$+/(X * 3) \div 10 \equiv 202.5$$

and the corresponding statistic is calculated as

$$m'_3 = \frac{1}{M} \sum_{i=1}^M d_i^3 \quad \text{or} \quad \frac{1}{M} \sum_{j=0}^9 j^3 f_j \quad (\text{Al.3-10})$$

or equivalently

$$F + . \times (0,19) * 3 .$$

Using the first and second moments, we can calculate the variance

$$\sigma^2 = \mu'_2 - \mu^2 = 28.5 - (4.5)^2 = 8.25 \quad (\text{Al.3-11})$$

or

$$(+/(X*2) \div 10) - (+/X \div 10) * 2 \equiv 8.25 ;$$

and employing the corresponding sample statistics, we can define the sample variance as

$$s^2 = m'_2 - m^2 \quad (\text{Al.3-12})$$

and the sample deviations as

$$s = \sqrt{s^2} .$$

The values of the above statistics for each of the three runs of TEST are recorded in Table Al-6.

TABLE Al-6

Frequency Test Statistics

Statistics	Theoretical Value	Observed Value		
		10000 Random Nos.	50000 Random Nos.	100000 Random Nos.
m	4.5	4.516	4.501	4.501
m_2'	28.5	28.666	28.462	28.501
m_3'	202.5	204.030	201.904	202.357
s^2	8.25	8.272	8.202	8.240
s	2.872	2.876	2.864	2.870

Recall that Tables Al-2, Al-3, and Al-4 contain the observed frequency of each digit, and notice that by using (Al.3-3) and the number of points generated we can calculate the expected frequency of each digit. Using the observed and expected frequencies, we can apply the chi-square test to determine if the numbers observed come from the population defined by (Al.3-3). We define a statistic $\hat{\chi}^2$ (CHI) such that

$$\hat{\chi}^2 = \sum_{j=0}^9 \frac{(f_j - M p(j))^2}{M p(j)} \quad (\text{Al.3-13})$$

$$= \sum_{j=0}^9 \frac{(f_j - M/10)^2}{M/10}$$

or

$$CHI \leftarrow +/((F - M \div 10) * 2) \times 10 \div M$$

where f_j is the observed frequency and $M p(j)$ is the expected frequency of the digit j . The statistic $\hat{\chi}^2$ tends asymptotically to a chi-square variate with 9 degrees of freedom as $M \rightarrow +\infty$. Thus we can assume that $\hat{\chi}^2$ has the probability density function

$$f(x) = \frac{1}{\Gamma(9/2) 2^{9/2}} x^{9/2-1} e^{-x/2} . \quad (\text{Al.3-14})$$

Therefore if we predetermine an $\alpha\%$ confidence level such that

$$\alpha/100 = \int_0^{c_o} f(x) dx , \quad (\text{Al.3-15})$$

then we will accept the hypothesis if $\hat{\chi}^2 \leq c_o$, otherwise we reject it at the $\alpha\%$ confidence level. For example, if $\alpha=90$ then $c_o=14.7$ and we would accept the hypothesis if $\hat{\chi}^2 \leq 14.7$, otherwise we reject the hypothesis. The

following Table A1-7 shows the results obtained by TEST. Notice the only questionable result is the last run.

TABLE A1-7

Chi-Square Statistic For The Frequency Test

Number of Digits generated	Critical Value c_0	Chi-Square Statistic $\hat{\chi}^2$
10000	14.7	6.632
50000	14.7	8.367
100000	14.7	14.97

b) Odd and Even Test

If the numbers generated are random, then there is an equal chance of generating an even or an odd digit. In this test the number of odd digits θ , and the number of even digits e , in the sample D are calculated as

$$e = \sum_{i=0}^4 f_{2i} \quad \text{and} \quad \theta = \sum_{i=1}^5 f_{2i-1} \quad (\text{A1.3-16})$$

or

$$E \leftarrow +/F[1 3 5 7 9] \quad \text{and}$$

$$OH \leftarrow +/F[2 4 6 8 10] .$$

We can then postulate that if the numbers in D are random, then on the average $M/2$ digits should be odd and $M/2$ digits should be even. Thus we can again test this hypothesis using the chi-square test where the statistic $\hat{\chi}^2$ (CHI) is defined as

$$\hat{\chi}^2 = \frac{2}{M} \{ (e - M/2)^2 + (o - M/2)^2 \} \quad (\text{A1.3-17})$$

or

$$CHI \leftarrow (2/M) \times ((E-M/2) * 2) + (O - M/2) * 2 .$$

This statistic also has an asymptotical chi-square distribution with one degree of freedom. Thus one test of the above hypothesis using a confidence level of 90% is to accept the numbers as random if $\hat{\chi}^2 \leq 2.71$, otherwise reject them as not being random. The $\hat{\chi}^2$ results of this test are shown in Table A1-8 and are all acceptable at the 90% level.

TABLE A1-8

Chi-Square Statistic For The Odd And Even Test

Number of Digits generated	Critical Value c_o	Chi-Square Statistic $\hat{\chi}^2$
10000	2.71	.0144
50000	2.71	.288
100000	2.71	.036

The output of this test consists of the observed relative frequency of odd and even digits along with the value of the appropriate chi-square statistic. This output is generated by statements 67 through 72 of *TEST*.

c) Poker Test

This test is concerned with the distribution of certain quadruples of these generated numbers, and hence is given the name "poker test". We restructure D into $M \div 4$ sets of quadruples and then test these quadruples. We are interested in particular combinations that may occur in these quadruples. Specifically, we want to know in how many of the quadruples there are all different digits, one pair of similar digits, two pairs of similar digits, three digits alike, and all four digits alike. For simplicity, let us call the event of all different digits event one, the event of one pair of digits event two, and so on.

The major problem in this test is finding an algorithm which will map the quadruples into the specific combinations using as few comparisons or tests as possible. This is necessary in order to economize on time and space. The algorithm used makes six comparisons which, if we represent a quadruple as (A, B, C, D), consist of the following tests:

$$A = B, \quad A = C, \quad A = D, \quad B = C, \quad B = D, \quad C = D. \quad (\text{Al.3-18})$$

Thus we construct a vector v_i such that

$$v_i = (A=B), (A=C), (A=D), (B=C), (B=D), C=D \quad (Al.3-19)$$

$$= (v_{ij}), \quad i = 1, 2, \dots, M \div 4 \\ j = 1, 2, \dots, 6$$

or

$$V \leftarrow (6, M \div 4) \rho (D[,1] = D[,2], (D[,1]=D[,3]),$$

$$(D[,1]=D[,4]), (D[,2]=D[,3]), (D[,2]=D[,4]),$$

$$D[,3] = D[,4],$$

where each integral quadruple is transformed into a logical 6-tuple. Then each of these 6-tuples is then summed giving us the following mapping:

$$\sum_{j=1}^6 v_{ij} = \begin{array}{ll} 0 & \text{none alike} \\ 1 & \text{one pair} \\ 2 & \text{two pairs} \\ 3 & \text{three of a kind} \\ 6 & \text{four of a kind} \end{array}, \quad (Al.3-20)$$

$i=1, 2, \dots, M \div 4$

or simply in APL

$$+/[1] V .$$

Hence, we have transformed the $M \div 4$ quadruples into a $M \div 4$ component vector U such that

$$U = (u_i) \text{ where } u_i = \sum_{j=1}^6 v_{ij}, \quad i=1, 2, \dots, M \div 4 \quad (\text{Al.3-21})$$

or

From the vector U we can obtain a frequency vector R where

$$R = (r_i), \quad i=1, 2, \dots, 5 \quad (\text{Al.3-22})$$

or

$$R \leftarrow +/(0 1 2 3 6) \circ. = U$$

and r_i is the number of times the i^{th} event occurred. We now have the observed frequency for each of these specific combinations. Let the vector S be defined as

$$S = (s_i) = (.504, .432, .027, .036, .001) \quad (\text{Al.3-23})$$

or

$$S \leftarrow 0.504 \quad 0.432 \quad 0.027 \quad 0.036 \quad 0.001$$

where s_i is the probability of the i^{th} event. Thus the expected frequency of the i^{th} event is $(M \div 4) \times s_i$.

Using this information, we can apply the chi-square test for goodness of fit. The statistic $\hat{\chi}^2$ (CHI) is defined as

$$\hat{\chi}^2 = \sum_{i=1}^5 \frac{4}{Ms_i} (r_i - \frac{s_i M}{4})^2 \quad (\text{Al.3-24})$$

or

$$CHI \leftarrow +/((R - S \times M \div 4) * 2) \div (S \times M) \div 4$$

which is approximately a chi-square variate with 4 degrees of freedom. Thus we can use this statistic to test the hypothesis, that the numbers are random, by accepting the hypothesis if $\hat{\chi}^2 < 7.78$ (at the 90% level) and rejecting it otherwise. The results produced were all acceptable and are shown in Table Al-9.

TABLE Al-9

Chi-Square Statistic For The Poker Test

Number of Digits generated .	Critical Value c_0	Chi-Square Statistic $\hat{\chi}^2$
10000	7.78	4.361
50000	7.78	1.044
100000	7.78	1.095

The output of this test consists of a table of the specific events, their theoretical probabilities, the observed frequencies, and the observed relative frequencies

and also the value of the χ^2 chi-square statistic. The accumulation of these data is done by TEST in statements 10 through 13 while the report is generated by statements 75 through 83.

d) Gap Test

Not only is the number of occurrences of any digit important, but in some simulations the number of digits between successive occurrences of this digit is also important. The number of digits between successive occurrences of a particular digit plus one is defined to be its gap length. For example, if the string of random numbers is

$$\begin{array}{r} 3 \ 5 \ 2 \ 8 \ 4 \ 6 \ 7 \ 3 \ 2 \ 5 \ 4 \ 5 \ 5 \\ \hline 8 \qquad \qquad \qquad 2 \ 1 \qquad \qquad \qquad 5 \end{array}$$

then the gap lengths for the digit 5 are 8, 2, 1, and 5. Since the probability of occurrence of any particular digit is 0.1, assuming the numbers are random, the probability of a gap length x is

$$p(x) = 0.1 (0.9)^{x-1}, \quad x=1, 2, \dots . \quad (\text{Al.3-25})$$

The distribution of the gap lengths is a geometric distribution with a probability of success of $p = 0.1$ such that the mean is $\mu=10$ and the variance is $\sigma^2=90$.

Given a string $D = (d_i)$, $i=1,2,\dots,M$, of random numbers, how do we find the gap lengths for each digit j , $j=0,1,\dots,9$? First, construct a vector T_j such that

$$T_j = (t_{jk}), \quad k=1,2,\dots, f_j \quad (\text{Al.3-26})$$

or

$$TJ \leftarrow (J = D) \times 1M$$

for each $j=0,1,\dots,9$ where t_{jk} indicates the position of the k^{th} occurrence of the digit j in the vector D . From this vector T_j we obtain a vector G_j of gap lengths for each j such that

$$G_j = (g_{jk}) = (t_{jk+1} - t_{jk}), \quad k=1,2,\dots, f_j - 1 \quad (\text{Al.3-27})$$

or

$$GJ \leftarrow (1 + TJ) - 1 + TJ.$$

The vector G_j contains the observed gap lengths from which we are able to obtain the observed frequencies and statistics of the gap lengths for the j^{th} digit.

The function TEST produces two relative frequency tables for the gap lengths. Two 23 by 7 tables are produced in which the first column contains the gap length (for gap lengths 1 to 23), the second column

contains the theoretical probability, and the remaining five columns contain the observed relative frequencies of the digits 0 to 4 in the first table and the digits 5 to 9 in the second table. The function also gives the number of gaps, mean, variance, standard deviation, maximum gap length, probability of the maximum gap length and the chi-square statistic for each digit.

The chi-square statistic is calculated on the basis of the lengths 1 to 23 ignoring any gap length greater than 23. Because the amount of information that can be stored is limited, the frequency table is truncated at 23. The gap lengths greater than 23 could have been pooled, but this would only permit another degree of freedom and at the same time introduce pooling errors. Hence there is not much to choose between the pooling method and the one employed here. Since the chi-square statistic is based on 23 classes, it has 22 degrees of freedom. If the hypothesis, that the observed data approximate the geometric distribution between 1 and 23, is to be accepted the $\hat{\chi}^2$ value should be less than 30.8, otherwise the hypothesis is to be rejected. The values of the chi-square statistics for each digit are shown in Table Al-10. Notice there are two rejections in the 30 tests.

TABLE A1-10Chi-Square Statistic For The Gap Test

Digit	Critical Value c_0	Number of Digits Generated		
		10000	50000	100000
0	30.8	13.7	20.2	24.0
1	30.8	14.4	29.6	33.0
2	30.8	14.6	10.1	15.7
3	30.8	20.5	22.2	31.6
4	30.8	19.3	18.4	16.7
5	30.8	18.5	23.0	21.8
6	30.8	27.1	17.4	10.8
7	30.8	21.7	23.3	18.7
8	30.8	18.1	28.4	23.1
9	30.8	21.3	22.6	28.2

The gap lengths are calculated by statements 14 to 22 of TEST while the report is generated by statements 86 to 107.

e) Coupon Collector Test

The coupon collector test calculates the length of the string of random numbers which is necessary in order for all ten digits to appear in the string. Random numbers are collected until all ten different digits have

been selected. The theoretical distribution of these set sizes, x , is

$$p(x) = \frac{1}{10^{x-1}} \sum_{i=0}^8 (-1)^i \binom{9}{i} (9-i)^{x-1}, \quad (\text{Al.3-28})$$

$$x = 10, 11, \dots$$

and has a mean of 29.290 and variance of 125.687.

Here again the main problem is to find a satisfactory algorithm which will find the coupon set size in a string D of random numbers. The method used is based on the recursive formula

$$z_0 = 0 \quad (\text{Al.3-29})$$

$$z_{k+1} = z_k + \max w_j^{(k)}$$

for $k=0,1,2,\dots, M-10$ or until we can no longer form the vector $w^{(k)}$. The vector $w^{(k)} = (w_j^{(k)})$, $j=0,1,\dots,9$ is defined such that $w_j^{(k)}$ is the index of the first occurrence of j in the vector

$$D^{(k)} = (d_{z_{K+1}}, d_{z_{K+2}}, \dots, d_M). \quad (\text{Al.3-30})$$

Notice, the vector $w^{(k)}$ can no longer be constructed when there exists a j which is not in $D^{(k)}$. If we define a vector

$$z = (z_{k+1} - z_k), \quad k=1,2,\dots, v, \quad (\text{Al.3-31})$$

then this vector contains the coupon set sizes. Z is used to obtain the frequency table and the statistics.

The results of this test is a 25 by 4 table in which the first column contains set sizes, the second column contains the observed frequencies, the third column contains the observed relative frequencies, and the fourth column contains the theoretical probabilities. Also shown are the number of completed sets, the mean set size, variance, standard deviation, maximum set size, probability of the maximum set size, and the value of the chi-square statistic. It is necessary in TEST to set the $\Pr(X=x) = 0$ for $x \geq 80$. The above results are generated by statements 110 to 122 of TEST.

The chi-square statistic is calculated from the truncated distribution ($x = 10, 11, \dots, 34$). However, the use of a chi-square test here must be viewed very critically. The reason for this is that $\hat{\chi}^2$ where

$$\hat{\chi}^2 = \sum \frac{(\text{observed frequency} - \text{expected frequency})^2}{\text{expected frequency}}$$

(Al.3-32)

only tends to a chi-square distribution. The rate at which χ^2 tends to a chi-square variate depends on the theoretical distribution involved as well as on the number of points observed. The original distribution (Al.3-28) must be well

behaved. Many rules regarding the theoretical distribution have been evolved by Cochran((1952)and (1954)) and Keeping (1962). For the distribution (Al.3-28), a unimodal distribution with well behaved tails, Cochran suggests that each frequency class has an expected frequency of at least one. In the coupon distribution $\Pr(X=10) \approx 3.6 \times 10^{-4}$, and the expected number of completed sets is $M \div 29.290$. If M is 10000 or 50000 the expected number of completed sets is 341 or 1707, respectively. Both of these values yield an expected frequency of less than one for the frequency class $X=10$. Almost 100000 random numbers are necessary to satisfy the above criterion so some compromise between the theoretical and practical value of this test must be made. The critical value of the chi-square statistic is 33.2 and resulting values of $\hat{\chi}^2$ are in Table Al-11.

TABLE Al-11Statistics For The Coupon Test

Statistic	Theore-tical Value	Number of Digits Generated : $N \times M$		
		10000	50000	100000
No. of sets .	(.03414) $N \times M$	343	1712	3407
Mean	29.290	29.149	29.204	29.327
Variance	125.687	111.013	122.090	124.604
St. de- viation	11.21	10.536	11.049	111.163
Max.set		76	86	94
Chi-square	33.2	15.91	31.41	28.30

f) Serial Test

If the numbers generated are random then the successor of a particular number is as likely to be one digit as another, i.e., the successor of a particular digit may be any one of the ten digits - all with equal probability. In order to make this test, first construct a sequence of pairs from the string of random numbers D by forming the pairs

$$(d_i, d_{i+1}), \quad i = 1, 2, \dots, M-1 \quad (\text{Al.3-33})$$

and (d_M, d_1)

or

$$D \leftarrow (2, M) \rho D, \quad 1 \notin D .$$

Then construct a two-way frequency table

$$F = (f_{ij}), \quad i, j = 0, 1, \dots, 9 \quad (\text{Al.3-34})$$

or

$$F \leftarrow +/(0, 199) \circ. = (1, 10) +. \times D$$

where f_{ij} indicates the number of times the digit j followed the digit i . Also define $f_{i.}$ as

$$f_{i.} = \sum_{j=0}^9 f_{ij} \quad (\text{Al.3-35})$$

or

$$FI \leftarrow +/F[I;],$$

and then calculate the following chi-square statistics

$$\hat{\chi}_1^2 = \frac{10}{M} \sum_{i=0}^9 (f_{i.} - \frac{M}{10})^2 \quad (\text{Al.3-36})$$

or

$$CHI\Delta 1 \leftarrow (+/((+/F) - M \div 10)*2) \times 10 \div M$$

and

$$\hat{\chi}_2^2 = \frac{100}{M} \sum_{i=0}^9 \sum_{j=0}^9 (f_{ij} - \frac{M}{100})^2 \quad (\text{Al.3-37})$$

or

$$CHI\Delta 2 \leftarrow ((+/+(F - M \div 100)*2) \times 100 \div M ,$$

where $\hat{\chi}_1^2$ will have 9 degrees of freedom while $\hat{\chi}_2^2$ has 99 degrees of freedom. It can be shown that $\hat{\chi}_3^2 = \hat{\chi}_2^2 - \hat{\chi}_1^2$ is asymptotically chi-square with $(10)^2 - 10 = 90$ degrees of freedom and $\hat{\chi}_4^2 = \hat{\chi}_2^2 - 2\hat{\chi}_1^2$ is also asymptotically chi-square with $(10-1)^2 = 81$ degrees of freedom (see Hull and Dobell (1962)).

The calculations for this test are done in statements 38 to 41, and the report which consists of the two-way frequency table F and the form values of the chi-square statistics is prepared in statements 125 to 136 of TEST. Table Al-12 contains the results of this test for the three trials.

TABLE A1-12Chi-Square Statistics For The Serial Test

Chi-Square Statistics	Critical Value c ₀ (90% level)	Number of Digits Generated		
		10000	50000	100000
$\hat{\chi}_1^2$	14.7	6.632	8.3672	14.97
$\hat{\chi}_2^2$	117.3	74.76	102.176	100.616
$\hat{\chi}_3^2$	107.6	68.128	93.8088	85.646
$\hat{\chi}_4^2$	97.6	61.496	85.4416	70.676

A1.4 Conclusions

The following pages contain the listing of TEST and an example of the function's output format. The output format was "cooked" slightly to produce a suitable output (i.e., to fit an 8½ by 11 page).

Generally, the numbers in the vector $~1+?M\rho 10$ (for $M \leftarrow 10000$ or 50000 or 100000) appear to have random number properties. However, it should be noted that the "randomness" of any particular set of such numbers cannot be guaranteed. One set of numbers may excel in the frequency test but may fail the gap test, and thus the purpose of the simulation will determine if the numbers are usable. Some numbers are

very poor random numbers and should not be used in any simulation experiment. The only way to be positive, that the random numbers are in fact random, is to test them, and therefore all simulation experiments should contain an analysis of the random numbers used.


```

VTEST[]▽
▽ N TEST M;A1;C1;DT;D1;D2;D3;D4;D5;D6;EXIT;E1;E2;E3;E4;E5;E6
;F1;F2;F3;I;K;L;R;S1;S3;S4;S5;S6;S7;S8;TM
[1] S4← 23 10 ρ, S5← 4 10 ρ S8←100ρ S7←4ρ S6←25ρ D3←10ρ S3←
5ρ S1←10ρ EXIT←0
[2] TM←, I21
[3] DT← 1+?Mρ 10
[4] F2←DT[I←1], E3←10
[5] REPEAT:F3←DT[1], E6←10
[6] D4← 23 10 ρ, D5← 4 10 ρ F1←100ρ E4←25ρ E5←5ρ D6←10ρ C1←
5ρ A1←10ρ 0
[7] WHICH:TM[ρ TM]←((I21)-TM[ρ TM])÷60
[8] A1←A1+K←+/ (0,19)◦.=DT
[9] S1←S1+K
[10] DT←((M÷4),4)ρ DT
[11] K←(DT[ ;1]=DT[ ;2]),(DT[ ;1]=DT[ ;3]),(DT[ ;1]=DT[ ;4]),(DT[ ;
2]=DT[ ;3]),(DT[ ;2]=DT[ ;4]),DT[ ;3]=DT[ ;4]
[12] C1←C1+K←+/ 6 3 2 1 0 ◦.=+/[1](6,M÷4)ρ K
[13] S3←S3+K
[14] D1←0
[15] DT←,DT
[16] HERE:→(M<DT,1,D1)/THERE
[17] D4[ ;D1+1]←D4[ ;D1+1]++/(123)◦.=D2←(D6[D1+1]=0)+L-D6[D1+1],
-1↓L←(L/(L←DT=D1)×1M)+M×PRINT|I-1
[18] D5[ ;D1+1]←(D5[1;D1+1]++/D2),(D5[2;D1+1]++/D2×2),(D5[
3;D1+1][ /D2),D5[4;D1+1]+ρ D2
[19] D6[D1+1]←L[ρ L]
[20] S4[ ;D1+1]←S4[ ;D1+1]++/(123)◦.=D2←(D3[D1+1]=0)+L-D3[D1+1],
-1↓L←(L/(L←DT=D1)×1M)+M×I-1
[21] S5[ ;D1+1]←(S5[1;D1+1]++/D2),(S5[2;D1+1]++/D2×2),(S5[
3;D1+1][ /D2),(ρ D2)+S5[4;D1+1]
[22] D3[D1+1]←L[ρ L]
[23] THERE:→(10≥D1←D1+1)/HERE
[24] K←E6,DT,E1←10
[25] THIS:K←((ρ K)L L←[ /K10,19)+K
[26] E1←E1,(0≠ρ K)/L
[27] →(0≠ρ K)/THIS
[28] E6←(E6,DT)[(+/E1)+1(M+ρ E6)-+/E1]
[29] K←E3,DT,E2←10
[30] THAT:K←((ρ K)L L←[ /K10,19)+K
[31] E2←E2,(0≠ρ K)/L
[32] →(0≠ρ K)/THAT
[33] E3←(E3,DT)[(+/E2)+1(M+ρ E3)-+/E2]
[34] E4←E4++/(9+125)◦.=E1
[35] E5←(E5[1]++/E1),(E5[2]++/E1×2),(E5[3][ /E1),E5[4]+ρ E1
[36] S6←S6++/(9+125)◦.=E2
[37] S7←(S7[1]++/E2),(S7[2]++/E2×2),(S7[3][ /E2),S7[4]+ρ E2
[38] K←0
[39] IT:F1[L+1]←F1[L+1]++/(L←(K×10)+0,19)◦.=(-0≠PRINT|I)+1↓(
10×F3,DT)+1ΦF3,DT
[40] S8[L+1]←S8[L+1]++/L◦.=(-I≠H)+1↓(10×F2,DT)+1ΦF2,DT

```



```

[41] →(9≥K←K+1)/IT
[42] F2←F2[1],DT[M]
[43] F3←F3[1],DT[M]
[44] →(0=PRINT|I)/WHAT
[45] I←I+1
[46] TM←TM,I21
[47] DT←-1+?Mρ10
[48] →WHICH
[49] WHAT:DT←10
[50] ('TEST NO. ',I÷PRINT,' FOR M= ',M÷PRINT×M,', RANDOM NUMBERS')
[51] (R←5ρ' ','TIME TAKEN WAS ',~1↑TM←((-PRINT)+TM),+/-(-PRINT)+TM;,' SECONDS')
[52] SMM: ' '
[53] ' '
[54] ' '
[55] (R;'FREQUENCY TEST FOR 10 CLASSES' )
[56] Q 4 10 ρ(0,19),A1,(A1÷M),((10)◦.≥10)+.×A1÷M
[57] ' '
[58] (R;'AVERAGE IS : ',L←(A1+.×0,19)÷M)
[59] (R;'SECOND MOMENT IS : ',K←(A1+.×(0,19)*2)÷M)
[60] (R;'THIRD MOMENT IS : ',(A1+.×(0,19)*3)÷M)
[61] (R;'VARIANCE IS : ',L←K-L*2)
[62] (R;'STANDARD DEVIATION IS : ',L*0.5)
[63] (R;'CHI-SQUARE STATISTIC : ',(+/(A1-M÷10)*2)×10÷M)
[64] (R;'WITH 9 DEGREES OF FREEDOM' )
[65] ' '
[66] ' '
[67] (R;'ODD AND EVEN TEST ')
[68] ' '
[69] (R;'RELATIVE FREQUENCY OF ODD IS : ',(L←+/A1[2 4 6 8 10])÷M)
[70] (R;'RELATIVE FREQUENCY OF EVEN IS : ',(K←+/A1[1 3 5 7 9])÷M)
[71] (R;'CHI-SQUARE STATISTIC IS : ',(+/((L,K)-M÷2)*2)×2÷M)
[72] (R;'WITH 1 DEGREE OF FREEDOM' )
[73] ' '
[74] ' '
[75] (R;'POKER TEST ON GROUPS OF FOUR' )
[76] ' '
[77] (R;'4 OF A KIND :XXXX P=.0010 ',C1[1];R;C1[1]÷M÷4)
[78] (R;'3 OF A KIND :XXYY P=.0360 ',C1[2];R;C1[2]÷M÷4)
[79] (R;'2 PAIRS :XXYY P=.0270 ',C1[3];R;C1[3]÷M÷4)
[80] (R;'1 PAIR :XXYZ P=.4320 ',C1[4];R;C1[4]÷M÷4)

```



```

[81] (R; 'BUST' :WXYZ P=.5040 ' ;C1[5];R;C1[5]÷M÷4)
[82] (R; 'CHI-SQUARE STATISTIC IS' :';+/((L-C1)*2)÷L←(0.001 0.036 0.027 0.432 0.504)×M÷4)
[83] (R; 'WITH 4 DEGREES OF FREEDOM')
[84] '
[85] '
[86] (R; 'GAP TEST')
[87] '
[88] (R; 'FREQUENCY TABLE FOR ALL DIGITS 0-4')
[89] (R; 'COL. 1 :GAP LENGTHS')
[90] (R; 'COL. 2 :PROBABILITY OF GAP LENGTHS')
[91] (R; 'COL. 3-7 :RELATIVE FREQUENCY OF GAPS FOR DIGITS 0-4')
[92] Ⓛ 7 23 Ⓜ(123),4 RND(K←0.1×0.9*^-1+123),,QD4[,;15]÷23 5 ⓂD5[4;15]
[93] '
[94] (R; 'FREQUENCY TABLE FOR ALL DIGITS 5-9')
[95] (R; 'COL. 1 :GAP LENGTHS')
[96] (R; 'COL. 2 :PROBABILITY OF GAP LENGTHS')
[97] (R; 'COL. 3-7 :RELATIVE FREQUENCY OF GAPS FOR DIGITS 5-9')
[98] Ⓛ 7 23 Ⓜ(123),4 RND K,,QD4[,5+15]÷23 5 ⓂD5[4;5+15]
[99] '
[100] (R; 'NUMBER OF GAPS IS' :';D5[4;])
[101] (R; 'MEAN IS' :';L←D5[1;]÷D5[4;])
[102] (R; 'VARIANCE IS' :';L←(D5[2;]÷D5[4;])-L*2)
[103] (R; 'STANDARD DEVIATION IS' :';L*0.5)
[104] (R; 'MAXIMUM GAP LENGTH IS' :';D5[3;])
[105] (R; 'PROBABILITY OF MAX GAP' :';0.1×0.9*^-1+D5[3;])
[106] (R; 'CHI-SQUARE STATISTIC IS' :';+/[(1)((K-D4)*2)÷K←(Q10 23 ⓂK)×23 10 ⓂD5[4;]])
[107] (R; 'WITH 22 DEGREES OF FREEDOM')
[108] '
[109] '
[110] (R; 'COUPON TEST')
[111] '
[112] (R; 'FREQUENCY TABLE OF SET SIZES')
[113] 5 RND Ⓛ 4 25 ⓂL,E4,(E4÷E5[4]),K←(0.1*L-1)×((-1*0,18)×(0,18)!9)+.×(9-0,18)○.*^-1+L←9+125
[114] '
[115] (R; 'NUMBER OF SETS' :';E5[4])
[116] (R; 'MEAN IS' :';L←E5[1]÷E5[4])
[117] (R; 'VARIANCE IS' :';L←(E5[2]÷E5[4])-L*2)
[118] (R; 'STANDARD DEVIATION IS' :';L*0.5)
[119] (R; 'MAXIMUM SET SIZE IS' :';L←E5[3])

```



```

[120] (R; 'PROBABALITY MAX SET      : ' ; KK×(
    0.1×L-1)×+/( -1×0, 18)×((0, 18)!9)×(9-0, 18)×-1+L←(KK←L≤
    80)×L)
[121] (R; 'CHI-SQUARE STATISTIC IS : ' ; +/((K-E4)*2)÷K←K×E5[
    4])
[122] (R; 'WITH 24 DEGREES OF FREEDOM')
[123] ' '
[124] ' '
[125] (R; 'SERIAL TEST')
[126] (R; 'TWO WAY FREQUENCY TABLE FOR (I,J)')
[127] ⓧ←F1← 10 10 ⓧF1
[128] ' '
[129] (R; 'CHI-SQUARE STATISTIC NO. 1: ' ; L←(10÷M)×+/( (+/F1)-M÷
    10)*2)
[130] (R; 'WITH 9 DEGREES OF FREEDOM')
[131] (R; 'CHI-SQUARE STATISTIC NO. 2: ' ; K←(100÷M)×+/+/(F1-M÷
    100)*2)
[132] (R; 'WITH 99 DEGREES OF FREEDOM')
[133] (R; 'CHI-SQUARE STATISTIC NO. 3: ' ; K-L)
[134] (R; 'WITH 90 DEGREES OF FREEDOM')
[135] (R; 'CHI-SQUARE STATISTIC NO. 4: ' ; K-2×L)
[136] (R; 'WITH 81 DEGREES OF FREEDOM')
[137] ' '
[138] ' '
[139] ' '
[140] ' '
[141] →(EXIT=1)/0
[142] →(N<I←I+1)/CONCLUSION
[143] M←M÷PRINT
[144] TM←TM, I21
[145] DT←-1+?Mρ10
[146] →REPEAT
[147] CONCLUSION: M←N×M÷PRINT
[148] K←(A1←S1), (C1←S3), (E4←S6), (E5←S7), (F1←S8), EXIT←1
[149] K←10
[150] D4←S4
[151] D5←S5
[152] ('GENERAL SUMMARY FOR ' ; M; ' POINTS')
[153] (R; 'TOTAL GENERATION TIME ' ; +/TM; ' SECONDS')
[154] (R; 'AVERAGE GENERATION TIME PER TRIAL ' ; (+/TM)÷n; ' SECONDS'
    )
[155] (R; 'AVERAGE GENERATION TIME PER NO. GENERATED ' ; (+/TM)÷M; ' SECONDS')
[156] →SMM

```


1 TEST 1000

TEST NO. 1 FOR M= 1000 RANDOM NUMBERS
TIME TAKEN WAS 0.25 SECONDS

FREQUENCY TEST FOR 10 CLASSES

0	95	0.0950	0.0950
1	97	0.0970	0.1920
2	97	0.0970	0.2890
3	93	0.0930	0.3820
4	117	0.1170	0.4990
5	111	0.1110	0.6100
6	110	0.1100	0.7200
7	94	0.0940	0.8140
8	96	0.0960	0.9100
9	90	0.0900	1.0000

AVERAGE IS : 4.4890
 SECOND MOMENT IS : 27.9690
 THIRD MOMENT IS : 195.5110
 VARIANCE IS : 7.8179
 STANDARD DEVIATION IS : 2.7960
 CHI-SQUARE STATISTIC : 7.5400
 WITH 9 DEGREES OF FREEDOM

ODD AND EVEN TEST

RELATIVE FREQUENCY OF ODD IS : 0.4850
 RELATIVE FREQUENCY OF EVEN IS : 0.5150
 CHI-SQUARE STATISTIC IS : 0.9000
 WITH 1 DEGREE OF FREEDOM

POKER TEST ON GROUPS OF FOUR

4 OF A KIND : XXXX	P=.0010	0	0.0000
3 OF A KIND : XXXY	P=.0360	8	0.0320
2 PAIRS : XXYY	P=.0270	5	0.0200
1 PAIR : XXYZ	P=.4320	115	0.4600
BUST : WXYZ	P=.5040	122	0.4880
CHI-SQUARE STATISTIC IS	:	1.3955	
WITH 4 DEGREES OF FREEDOM			

GAP TEST

FREQUENCY TABLE FOR ALL DIGITS 0-4

COL. 1 :GAP LENGTHS

COL. 2 :PROBABALITY OF GAP LENGTHS

COL. 3-7 :RELATIVE FREQUENCY OF GAPS FOR DIGITS 0-4

1	0.1000	0.0851	0.0833	0.0938	0.0978	0.1207
2	0.0900	0.0957	0.1042	0.0729	0.0435	0.0948
3	0.0810	0.0745	0.1146	0.0625	0.0870	0.1293
4	0.0729	0.0851	0.0938	0.1250	0.0217	0.0603
5	0.0656	0.0532	0.0521	0.0833	0.0978	0.0517
6	0.0590	0.0319	0.0521	0.0625	0.0543	0.0690
7	0.0531	0.0851	0.0417	0.0208	0.0652	0.0259
8	0.0478	0.0426	0.0208	0.0521	0.0761	0.0517
9	0.0430	0.0106	0.0417	0.0521	0.0652	0.0776
10	0.0387	0.0319	0.0313	0.0104	0.0652	0.0603
11	0.0349	0.0319	0.0313	0.0104	0.0326	0.0345
12	0.0314	0.0851	0.0208	0.0417	0.0109	0.0172
13	0.0282	0.0532	0.0104	0.0521	0.0217	0.0345
14	0.0254	0.0532	0.0521	0.0104	0.0217	0.0086
15	0.0229	0.0106	0.0104	0.0313	0.0543	0.0172
16	0.0206	0.0106	0.0000	0.0000	0.0109	0.0086
17	0.0185	0.0000	0.0208	0.0313	0.0109	0.0086
18	0.0167	0.0106	0.0104	0.0208	0.0326	0.0086
19	0.0150	0.0106	0.0208	0.0000	0.0109	0.0086
20	0.0135	0.0106	0.0417	0.0104	0.0217	0.0172
21	0.0122	0.0000	0.0208	0.0104	0.0000	0.0172
22	0.0109	0.0000	0.0104	0.0000	0.0109	0.0172
23	0.0098	0.0106	0.0313	0.0313	0.0000	0.0086

COL. 3-7:RELATIVE FREQUENCY OF GAPS FOR DIGITS 5-9

1	0.1000	0.1091	0.1651	0.1183	0.1053	0.0674
2	0.0900	0.0818	0.1101	0.0538	0.0842	0.0674
3	0.0810	0.1000	0.1009	0.0968	0.0947	0.1236
4	0.0729	0.1273	0.0275	0.0430	0.0526	0.0674
5	0.0656	0.0727	0.0550	0.0753	0.0737	0.0787
6	0.0590	0.0455	0.0550	0.0430	0.0632	0.0449
7	0.0531	0.0818	0.0642	0.1290	0.0632	0.0449
8	0.0478	0.0727	0.0459	0.0215	0.0632	0.0225
9	0.0430	0.0364	0.0459	0.0323	0.0632	0.0337
10	0.0387	0.0000	0.0642	0.0323	0.0211	0.0112
11	0.0349	0.0364	0.0275	0.0215	0.0211	0.0562
12	0.0314	0.0182	0.0183	0.0538	0.0316	0.0225
13	0.0282	0.0182	0.0092	0.0215	0.0000	0.0449
14	0.0254	0.0182	0.0275	0.0000	0.0000	0.0337
15	0.0229	0.0364	0.0092	0.0215	0.0211	0.0337
16	0.0206	0.0091	0.0092	0.0215	0.0211	0.0449
17	0.0185	0.0182	0.0183	0.0108	0.0211	0.0112
18	0.0167	0.0091	0.0092	0.0000	0.0421	0.0000
19	0.0150	0.0273	0.0000	0.0108	0.0105	0.0337
20	0.0135	0.0000	0.0183	0.0108	0.0105	0.0112
21	0.0122	0.0000	0.0092	0.0103	0.0105	0.0112
22	0.0109	0.0000	0.0092	0.0430	0.0105	0.0112
23	0.0098	0.0000	0.0000	0.0108	0.0316	0.0000

X	GAP	MEAN	GAP STATISTICS		MAX	P(MAX)	CHI
			VAR	S.D.			
0	94	10.223	84.450	9.190	45	0.0010	25.11
1	96	10.292	95.165	9.755	54	0.0004	22.52
2	96	10.354	89.458	9.458	42	0.0013	23.91
3	92	10.685	108.151	10.400	58	0.0002	21.67
4	116	8.509	64.267	8.017	41	0.0015	16.27
5	110	9.009	104.482	10.222	66	0.0001	22.47
6	109	9.092	88.450	9.405	40	0.0016	18.19
7	93	10.495	94.228	9.707	53	0.0004	30.68
8	95	10.358	119.535	10.933	63	0.0001	17.36
9	89	11.135	103.487	10.173	53	0.0004	17.49

CHI-SQUARE STATISTIC HAS 22 D. OF F.

COUPON TEST

FREQUENCY TABLE OF SET SIZES

10	0	0.00000	0.00036
11	0	0.00000	0.00163
12	1	0.03030	0.00419
13	0	0.00000	0.00808
14	1	0.03030	0.01305
15	0	0.00000	0.01863
16	0	0.00000	0.02436
17	0	0.00000	0.02978
18	0	0.00000	0.03458
19	1	0.03030	0.03853
20	3	0.09091	0.04154
21	0	0.00000	0.04359
22	3	0.09091	0.04473
23	1	0.03030	0.04507
24	6	0.18182	0.04471
25	0	0.00000	0.04377
26	0	0.00000	0.04238
27	1	0.03030	0.04065
28	3	0.09091	0.03867
29	2	0.06061	0.03653
30	2	0.06061	0.03430
31	1	0.03030	0.03204
32	0	0.00000	0.02980
33	0	0.00000	0.02760
34	0	0.00000	0.02547

NUMBER OF SETS : 33.0000
 MEAN IS : 29.4545
 VARIANCE IS : 125.3994
 STANDARD DEVIATION IS : 11.1982
 MAXIMUM SET SIZE IS : 61.0000
 PROBABILITY MAX SET : 0.0018
 CHI-SQUARE STATISTIC IS : 38.2250
 WITH 24 DEGREES OF FREEDOM

SERIAL TEST
 TWO WAY FREQUENCY TABLE FOR (I, J)

8	15	6	7	14	15	9	9	4	8
12	8	6	9	9	12	10	11	4	16
12	12	9	9	10	6	11	10	10	8
11	8	15	9	9	13	10	3	9	6
11	11	12	13	14	15	9	10	12	10
8	8	9	9	15	12	12	12	18	8
12	11	7	12	11	6	18	13	11	9
7	5	6	8	14	11	11	11	10	11
8	10	12	8	9	13	9	9	10	8
6	9	15	9	12	8	11	6	8	6

CHI-SQUARE STATISTIC NO. 1: 7.5400

WITH 9 DEGREES OF FREEDOM

CHI-SQUARE STATISTIC NO. 2: 85.4000

WITH 99 DEGREES OF FREEDOM

CHI-SQUARE STATISTIC NO. 3: 77.8600

WITH 90 DEGREES OF FREEDOM

CHI-SQUARE STATISTIC NO. 4: 70.3200

WITH 81 DEGREES OF FREEDOM

APPENDIX 2

THE MONTE CARLO SIMULATION PROGRAM

In this Appendix we shall briefly discuss the Monte Carlo simulation of stochastic linear programs and describe how the following program may be used to perform this simulation. The Monte Carlo simulation program collects data on the solution of randomly generated linear programs. Using the data obtained from the simulation, we hope to be able to predict the behaviour of the stochastic linear programming model (see Sections 4.6 and 5.3).

The stochastic linear program can be represented as

$$z^* = \max_{\underline{x}} \underline{c}' \underline{x} \quad (\text{A2-1})$$

subject to

$$A \underline{x} \leq \underline{b} \quad (\text{A2-2})$$

and

$$\underline{x} \geq \underline{0} \quad , \quad (\text{A2-3})$$

where A is an m by n matrix, and \underline{b} , \underline{c} , and \underline{x} are dimensioned accordingly. The elements of A and \underline{b} are mutually independent normal variates where $E[A] = M_0$, $\text{Var}(A) = S$, $E[\underline{b}] = \underline{b}_0$, and $\text{Var}(\underline{b}) = V$. If the stochastic linear

program can be expressed in the above form, where M_o , S , b_o , v , and c are known constants, then the following program can be used to perform the Monte Carlo simulation.

The following table, Table A2-1, contains the input or data set specifications for the Monte Carlo simulation program.

TABLE A2-1

Data Set Specifications For The Monte Carlo Routine

Block	No. Cards	Variables	Columns	Type	Comment
1	1	NR	1-4	Integer	no. of rows in M_o
			5-8	Integer	no. of columns in M_o
2	$\lceil NR \times NC \div 8 \rceil$	EA	1-80	Real	elements of M_o in row order and F10.5 format
3	$\lceil NR \times NC \div 8 \rceil$	VA	1-80	Real	elements of S in row order and F10.5 format
4	$\lceil NR \div 8 \rceil$	EB	1-80	Real	elements of b_o in F10.5 format
5	$\lceil NR \div 8 \rceil$	VB	1-80	Real	elements of v in F10.5 format
6	$\lceil NC \div 8 \rceil$	C	1-80	Real	elements of c in F10.5 format
7	1	NUMBER	1-10	Integer	the no. of linear programs to be generated

The operator \lceil is the greatest integer operator.

This data set must define a stochastic linear program of the form outlined above.

After the program has read in the data set, it initializes the frequency tables for z^* and \underline{x}_i , $i=1,2,\dots,m+n$. These frequency tables depend upon the particular model being simulated, and therefore must be defined in relation to the data set. Using the subroutines BOX and RANDOM, the program generates random points and creates the corresponding linear programs. Statistics are kept for random points which are reported at the end of the simulation, so that the randomness of the points may be checked. The simplex algorithm (SIMPLX) is used to solve the linear programs or to discover cycling or the existence of unbounded solutions. If there exists an optimal feasible solution to the linear program, then the statistics for z^* and \underline{x} are collected. The simulation continues until the required number (NUMBER) of linear programs have been generated.

Once the required number of linear programs has been generated, the output routine begins. The output from this Monte Carlo simulation begins with the average-value linear program corresponding to (A2-1), (A2-2), and (A2-3), i.e.,

$$\max_{\underline{x}} \underline{c}' \underline{x} \quad (A2-4)$$

subject to

$$M_o \underline{x} \leq b_o \quad (A2-5)$$

and

$$\underline{x} \geq 0 \quad . \quad (A2-6)$$

Next the program reports the number of linear programs generated, the mean and variance of z^* , the mean of x_i , $i=1,2,\dots, m+n$, and the variance-covariance matrix of \underline{x} . A frequency table is printed containing the frequency of the pair of optimal basic variables (x_i, x_j) , $i \neq j$. Frequency tables for z^* and x_i , $i=1,2,\dots, m+n$, are also given and are of the same form as Tables 5-3, 5-4, 5-5, and 5-6. Lastly, the program reports the mean and variance-covariance matrix of the generated random points.

The following Monte Carlo simulation program requires the FORTRAN subroutines: TABLE, SIMPLX, and BOX. The subroutine TABLE is a frequency tabling routine which uses a binary search technique. The subroutine SIMPLX is the FORTRAN implementation of the simplex algorithm. The Box-Muller algorithm is used in BOX to transform uniform random variates into standard normal variates. This subroutine requires the uniform random number generator, RANDOM, which is the assembly language version of the APL random number generator.

This Monte Carlo simulation program can be used as a basis for Tintner's Sampling method where only the method of generating and weighting the points needs to be changed. The Monte Carlo program can also be modified to either simulate or sample systems of linear equations. The following pages contain the main line listing of the Monte Carlo program, and the listing of the subroutines RANDOM, SIMPLX, TABLE, and BOX. The results obtained by using the Monte Carlo program to simulate Simplex #2 follow the program listings.

C
C SIMULATION OR MONTE CARLO
C TECHNIQUE FOR SOLVING
C STOCHASTIC LINEAR PROGRAMS
C G. LINKS
C

```
REAL*8 TAIL(5,64,2)/640*0.0D0/, MEAN/0.0D0/, MX(4)/4*0.0D0/,  

1M2X(4,4) , X(4) , DT(3,5), EA(2,2), VA(2,2), EB(2), VB(2), C(2), DC(5),  

2DX(2), R(6), PTS(6)/6*0.0D0/, VPTS(6,6)/36*0.0D0/, SUM(64), M2  

INTEGER FREQ(4,4)/16*0/, INDEX(2), INT/16807/, KOUNT/0/, KORE/0/  

DATA M2X/16*0.0D0/, M2/0.0D0/
```

C
C READ DATA FOR PROBLEM
C

```
READ(5,200) NR,NC  

READ(5,201)((EA(I,J),J=1,NC),I=1,NR)  

READ(5,201)((VA(I,J),J=1,NC),I=1,NR)  

READ(5,201)(EB(I),I=1,NR)  

READ(5,201)(VB(I),I=1,NR)  

READ(5,201)(C(I),I=1,NC)  

READ(5,202) NUMBER
```

```
200 FORMAT(2I4)  

201 FORMAT(8F10.5)  

202 FORMAT(I10)
```

C
C TEST FOR STOP
C

```
DO 3 I=1,64  

TAIL(1,I,1)=16.0D0+DFLOAT(I-1)*0.125D0  

TAIL(2,I,1)=(-2.0D0)+DFLOAT(I-1)*0.0625D0  

TAIL(3,I,1)=8.0D0+DFLOAT(I-1)*0.0625D0  

TAIL(4,I,1)=0.5D0+DFLOAT(I-1)*0.125D0  

3   TAIL(5,I,1)=(-2.0D0)+DFLOAT(I-1)*0.0625D0  

5   KOUNT=KOUNT+1  

IF(KOUNT.GT.NUMBER) GO TO 100
```

C
C GENERATION OF POINT
C

```
CALL BOX((NR+NR*NC),R,INT)  

DO 75 I=1,NR  

DO 75 J=1,NC  

K=J+(I-1)*NR  

R(K)=EA(I,J)+R(K)*DSQRT(VA(I,J))  

75   DT(I,J)=R(K)  

DO 76 I=1,NR  

K=NR*NC+I  

R(K)=EB(I)+R(K)*DSQRT(VB(I))  

76   DT(I,NR+NC+1)=R(K)  

K=NR+NR*NC  

DO 77 I=1,K  

PTS(I)=PTS(I)+R(I)  

DO 77 J=1,K  

77   VPTS(I,J)=VPTS(I,J)+R(I)*R(J)
```

C
C SET UP NEW PROBLEM


```

C
K=NR+NC+1
DO 78 I=1,K
DC(I)=0.0D0
IF(I.LE.NC) DC(I)=C(I)
78 CONTINUE
DO 79 I=1,NR
INDEX(I)=NC+I
DO 80 I=1,NR
DO 80 J=1,NR
DT(I,NC+J)=0.0D0
IF(J.EQ.I) DT(I,NC+J)=1.0D0
80 CONTINUE
MR=NR+1
MC=NR+NC+1

C
C   SOLVE NEW PROBLEM
C
CALL SIMPLX(MR,MC,DT,DC,INDEX,L,ZMAX,DX)
IF(L.EQ.1) GO TO 50
KORE=KORE+1
IF(L.EQ.2) GO TO 25
WRITE(6,900) KOUNT
900 FORMAT(' ', 'PROBLEM NO. ',I10,' IS CYCLING')
GO TO 5
25 WRITE(6,901) KOUNT
901 FORMAT(' ', 'PROBLEM NO. ',I10,' HAS AN UNBOUNDED SOLUTION')
GO TO 5

C
C   DATA COLLECTION
C
50 MEAN=MEAN+ZMAX
M2=M2+ZMAX*ZMAX
K=NR+NC
DO 51 I=1,K
51 X(I)=0.0D0
DO 52 I=1,NR
X(INDEX(I))=DX(I)
FREQ(INDEX(1),INDEX(2))=FREQ(INDEX(1),INDEX(2))+1
DO 53 I=1,K
MX(I)=MX(I)+X(I)
DO 53 J=1,K
53 M2X(I,J)=M2X(I,J)+X(I)*X(J)

C
C   TABLE DATA
C
CALL TABLE(TAIL,64,ZMAX,1.0D0,1)
DO 60 I=2,5
CALL TABLE(TAIL,64,X(I-1),1.0D0,I)
60 CONTINUE
GO TO 5

C
C   OUTPUT ROUTINE
C
100 K=NR+NC

```



```
NUMBER=NUMBER-KORE  
MEAN=MEAN/DFLOAT(NUMBER)  
M2=(M2/DFLOAT(NUMBER))-MEAN*MEAN
```

170

```
DO 101 I=1,K  
101 MX(I)=MX(I)/DFLOAT(NUMBER)  
DO 102 I=1,K  
DO 102 J=1,K  
102 M2X(I,J)=(M2X(I,J)/DFLOAT(NUMBER))-MX(I)*MX(J)  
DO 30 I=1,3  
K=I+1  
DO 30 J=K,4  
30 FREQ(I,J-1)=FREQ(I,J)+FREQ(J,I)  
WRITE(6,902)  
902 FORMAT('1','MONTE CARLO METHOD FOR SOLVING'  
'0','THE STOCHASTIC LINEAR PROGRAM :')  
WRITE(6,903)(C(I),I,I=1,NC)  
903 FORMAT('0',5X,'MAXIMIZE Z, WHERE '0',5X,'Z = ',F12.5,' X',I1,  
1' + ',F12.5,' X',I1)  
WRITE(6,904)((FA(I,J),J,J=1,NC),EB(I),I=1,NR)  
904 FORMAT('0',5X,'SUBJECT TO'/(0',5X,1(F12.5,' X',I1,' + '),  
1F12.5,' X',I1,' .LE. ',F12.5))  
WRITE(6,905) NUMBER  
905 FORMAT('0','AND USING ',I10,' POINTS')  
WRITE(6,906) MEAN ,M2  
906 FORMAT('0','APPROX. VALUE OF E(Z) IS ',F12.5/'0','APPROX. VALUE  
1'OF VAR(Z) IS',F11.5)  
WRITE(6,907)(I, MX(I),I=1,4)  
907 FORMAT('0','APPROX. VALUE OF E(X',I1,') IS',F12.5)  
WRITE(6,908)((M2X(I,J),J=1,4),I=1,4)  
908 FORMAT('0','APPROX. VAR-COVARIANCE MATRIX IS'/(0',4(E12.5,2X))  
WRITE(6,909)  
909 FORMAT('0','FREQUENCY OF PAIRS MATRIX')  
WRITE(6,910)(J,J=2,4)  
910 FORMAT( '0',1X,3(14X,I1))  
WRITE(6,911)(I,(FREQ(I,J),J=I,3),I=1,3)  
911 FORMAT('0',I1,3(5X,I10)/'0',I1,15X,2(5X,I10)/'0',I1,35X,I10)  
DO 35 J=1,5  
TAIL(J,1,2)=TAIL(J,1,2)/DFLOAT(NUMBER)  
SUM(1)=TAIL(J,1,2)  
DO 31 I=2,64  
TAIL(J,I,2)=TAIL(J,I,2)/DFLOAT(NUMBER)  
31 SUM(I)=SUM(I-1)+TAIL(J,I,2)  
IF(J.NE.1) GO TO 32  
WRITE(6,912)  
912 FORMAT('1','THE APPROX. DISTRIBUTION OF Z')  
GO TO 33  
32 K=J-1  
WRITE(6,913) K  
913 FORMAT('1','THE APPROX. MARGINAL DISTRIBUTION OF X',I1)  
33 CONTINUE  
WRITE(6,914)  
914 FORMAT('0',5X,'T',10X,'R.FREQ.',7X,'C.R.FREQ.',18X,'T',10X,  
1'R.FREQ.',7X,'C.R.FREQ.')  
DO 34 I=1,32  
K=32+1
```



```

      WRITE(6,915)(TAIL(J,I,L),L=1,2),SUM(I),(TAIL(J,K,N),N=1,2),SUM(
915  FORMAT(' ',3(E12.5,2X),10X,3(E12.5,2X))
34   CONTINUE
      WRITE(6,916)
916  FORMAT('0'// '/'0','T          : THE UPPER BOUND OF AN OPEN-CLOSED
     1' INTERVAL')
      WRITE(6,917)
917  FORMAT('0','R.FREQ.    : RELATIVE FREQUENCY OF THE INTERVAL T')
      WRITE(6,918)
918  FORMAT('0','C.R.FREQ. : CUMULATIVE RELATIVE FREQUENCY UP TO & '
     1' INCLUDING THE INTERVAL T')
35   CONTINUE
      KOUNT=KOUNT-1
      K=NR+NR*NC
      DO 36 I=1,K
36   PTS(I)=PTS(I)/DFLOAT(KOUNT )
      DO 37 I=1,K
      DO 37 J=1,K
37   VPTS(I,J)=(VPTS(I,J)/DFLOAT(KOUNT ))-PTS(I)*PTS(J)
      WRITE(6,919)
919  FORMAT('1','SUMMARY OF THE RANDOM NUMBERS'// '/')
      WRITE(6,920)(I,PTS(I),I=1,K)
920  FORMAT('0','THE MEAN VALUE OF R(',I2,',') IS ',E10.3)
      WRITE(6,921)((VPTS(I,J),J=1,K),I=1,K)
921  FORMAT(' ','THE VARIANCE-COVARIANCE MATRIX IS'/(0',6(E10.3,2X))
      STOP
      END

```

STMT SOURCE STATEMENT

1	RANDOM	CSECT	
2		USING *,15	INITIAL LINKAGE
3		STM 2,5,28(13)	
4		LM 2,3,0(1)	LOAD ADDRESSES OF VARIABLE
5		L 5,A	COMPUTE NEXT INTEGER
6		M 4,0(2)	X(I+1)=AX(I)(MOD P)
7		D 4,P	
8		ST 4,0(2)	
9		SRL 4,7	COMPUTE NEXT REAL
10		A 4,CHAR	
11		ST 4,0(3)	
12		LM 2,5,28(13)	TERMINAL LINKAGE
13		MVI 12(13),X'FF'	RETURN
14		BR 14	
15	CHAR	DC F'1073741824'	CONSTANTS. CHAR FIRST
16	A	DC F'16807'	SO A IS ON DOUBLE WORD
17	P	DC F'2147483647'	BOUNDARY. MAKES LM
18		END ,	INSTRUCTION FASTER.


```

SUBROUTINE SIMPLX(M1,N1,A,C,BASIS,S,ZMAX,SOLN)
REAL*8 A(M1,N1),C(40),SOLN(20),CBASIS(20),ZMAX
INTEGER BASIS(20), STEP, S
COMMON STEP
M=M1-1
N=N1-1
ZERO=1.0D-6
STEP=0
DO 1 I=1,M
CBASIS(I)=C(BASIS(I))
1 CONTINUE
C(N1)=0.0D0
DO 2 J=1,N1
SUM=0.0D0
DO 3 I=1,M
SUM=SUM+A(I,J)*CBASIS(I)
3 CONTINUE
A(M1,J)=SUM-C(J)
2 CONTINUE
18 DO 4 I=1,M
SOLN(I)=A(I,N1)
4 CONTINUE
ZMAX=A(M1,N1)
K=0
XMIN=0.0D0
DO 8 J=1,N
X=A(M1,J)
IF (X.GE.-ZERO) GO TO 8
K=K+1
IF(X.GE.XMIN) GO TO 8
XMIN=A(M1,J)
KIN=J
8 CONTINUE
IF (K.GT.0) GO TO 9
S=1
RETURN
9 K=0
XMIN=1.0D+20
DO 10 I=1,M
X=A(I,KIN)
IF (X.LE.ZERO) GO TO 10
K=K+1
X=A(I,N1)/X
IF(X.GE.XMIN) GO TO 10
XMIN=X
KOUT=I
10 CONTINUE
IF(K.GT.0) GO TO 11
S=2
RETURN
11 K=BASIS(KOUT)
BASIS(KOUT)=KIN
CBASIS(KOUT)=C(KIN)
STEP=STEP+1
IF(STEP.LE.2*M) GO TO 15

```



```

S=3
RETURN
15 P=A(KOUT,KIN)
DO 13 J=1,N1
A(KOUT,J)=A(KOUT,J)/P
13 CONTINUE
DO 16 I=1,M1
IF (I.EQ.KOUT) GO TO 16
P=A(I,KIN)
DO 14 J=1,N1
A(I,J)=A(I,J)-P*A(KOUT,J)
14 CONTINUE
16 CONTINUE
GO TO 18
END

```

TABLE

```

SUBROUTINE TABLE(A,N,X,Y,J)
REAL*8 A(5,N,2),X,Y
I=N/2
M=I
1 IF(X.LE.A(J,M,1)) GO TO 3
IF(I.EQ.1) GO TO 2
I=I/2
M=M+I
GO TO 1
2 A(J,M+1,2)=A(J,M+1,2)+Y
RETURN
3 IF(I.NE.1) GO TO 4
A(J,M,2)=A(J,M,2)+Y
RETURN
4 I=I/2
M=M-I
5 IF(A(J,M,1).LT.X) GO TO 7
IF(I.EQ.1) GO TO 6
I=I/2
M=M-I
GO TO 5
6 A(J,M,2)=A(J,M,2)+Y
RETURN
7 IF(I.EQ.1) GO TO 8
I=I/2
M=M+I
GO TO 1
8 A(J,M+1,2)=A(J,M+1,2)+Y
RETURN
END

```

BOX

```

SUBROUTINE BOX(N,R,INT)
REAL*8 R(N),POINT(12),PI/3.141592653589793/,X/0.0D0/
K=N*2
DO 250 J=1,K
CALL RANDOM(INT,X)
250 POINT(J)=X
DO 25 I=1,N
25 R(I)=DCOS(2.0D0*PI*POINT(N+I))*(-2.0D0*DLOG(POINT(I)))**0.5
RETURN
END

```


MONTE CARLO METHOD FOR SOLVING

THE STOCHASTIC LINEAR PROGRAM :

MAXIMIZE Z, WHERE

$$Z = 1.00000 X_1 + 2.00000 X_2$$

SUBJECT TO

$$3.00000 X_1 + 1.00000 X_2 \leq 15.00000$$

$$1.00000 X_1 + 1.00000 X_2 \leq 10.00000$$

AND USING 19000 POINTS

APPROX. VALUE OF E(Z) IS 20.02208

APPROX. VALUE OF VAR(Z) IS 1.03841

APPROX. VALUE OF E(X1) IS 0.02531

APPROX. VALUE OF E(X2) IS 9.99839

APPROX. VALUE OF E(X3) IS 4.91831

APPROX. VALUE OF E(X4) IS 0.00002

APPROX. VAR-COVARIANCE MATRIX IS

0.51514D-01	-0.25795D-01	-0.12447D 00	-0.49000D-06
-0.25795D-01	0.27255D 00	-0.20215D 00	0.23693D-04
-0.12447D 00	-0.20215D 00	0.17002D 01	-0.95226D-04
-0.49000D-06	0.23693D-04	-0.95226D-04	0.71221D-05

FREQUENCY OF PAIRS MATRIX

	2	3	4
1	245	0	0
2		18754	1
3			0

THE APPROX. DISTRIBUTION OF Z

T	R.FREQ.	C.R.FREQ.	T	R.FREQ.	C.R.FREQ.	T	R.FREQ.	C.R.FREQ.
0.16000D	02	0.0	0.20000D	02	0.49537D 00	0.48684D-01	0.49537D 00	0.54437D 00
0.16125D	02	0.0	0.20125D	02	0.49000D-01	0.47158D-01	0.59153D 00	0.59153D 00
0.16250D	02	0.0	0.20250D	02	0.47158D-01	0.46947D-01	0.63847D 00	0.63847D 00
0.16375D	02	0.0	0.20375D	02	0.46947D-01	0.44474D-01	0.68295D 00	0.68295D 00
0.16500D	02	0.0	0.20500D	02	0.44474D-01	0.42105D-01	0.72626D 00	0.72626D 00
0.16625D	02	0.0	0.20625D	02	0.43316D-01	0.39105D-01	0.76537D 00	0.76537D 00
0.16750D	02	0.0	0.20750D	02	0.39105D-01	0.34368D-01	0.79974D 00	0.79974D 00
0.16875D	02	0.0	0.20875D	02	0.34368D-01	0.21000D 02	0.83284D 00	0.83284D 00
0.17000D	02	0.0	0.21000D	02	0.33105D-01	0.21125D 02	0.86105D 00	0.86105D 00
0.17125D	02	0.0	0.21125D	02	0.28211D-01	0.13684D-02	0.90516D 00	0.90516D 00
0.17250D	02	0.0	0.21250D	02	0.22368D-01	0.19474D-02	0.92316D 00	0.92316D 00
0.17375D	02	0.0	0.21375D	02	0.21737D-01	0.18947D-02	0.94005D 00	0.94005D 00
0.17500D	02	0.0	0.21500D	02	0.18000D-01	0.16895D-01	0.95247D 00	0.95247D 00
0.17625D	02	0.0	0.21625D	02	0.16895D-01	0.14947D-01	0.96279D 00	0.96279D 00
0.17750D	02	0.0	0.21750D	02	0.12421D-01	0.10421D-01	0.97026D 00	0.97026D 00
0.17875D	02	0.0	0.21875D	02	0.10316D-01	0.73158D-02	0.97753D 00	0.97753D 00
0.18000D	02	0.0	0.22000D	02	0.74737D-02	0.57895D-02	0.98405D 00	0.98405D 00
0.18125D	02	0.0	0.22125D	02	0.72632D-02	0.78421D-02	0.99158D 00	0.99158D 00
0.18250D	02	0.0	0.22250D	02	0.65263D-02	0.10053D-01	0.99332D 00	0.99332D 00
0.18375D	02	0.0	0.22375D	02	0.39474D-02	0.13632D-01	0.99470D-02	0.99470D-02
0.18500D	02	0.0	0.22500D	02	0.35789D-02	0.15053D-01	0.99521D 00	0.99521D 00
0.18625D	02	0.0	0.22625D	02	0.17368D-02	0.17211D-01	0.99653D 00	0.99653D 00
0.18750D	02	0.0	0.22750D	02	0.18947D-02	0.19737D-01	0.99763D 00	0.99763D 00
0.18875D	02	0.0	0.22875D	02	0.13158D-02	0.25053D-01	0.99826D 00	0.99826D 00
0.19000D	02	0.0	0.23000D	02	0.11053D-02	0.28842D-01	0.99900D 00	0.99900D 00
0.19125D	02	0.0	0.23125D	02	0.63158D-03	0.30947D-01	0.99953D 00	0.99953D 00
0.19250D	02	0.0	0.23250D	02	0.73684D-03	0.36421D-01	0.99979D 00	0.99979D 00
0.19375D	02	0.0	0.23375D	02	0.26316D-03	0.40211D-01	0.99926D 00	0.99926D 00
0.19500D	02	0.0	0.23500D	02	0.26316D-03	0.43474D-01	0.99989D 00	0.99989D 00
0.19625D	02	0.0	0.23625D	02	0.10526D-03	0.46632D-01	0.10526D-03	0.10526D-03
0.19750D	02	0.0	0.23750D	02	0.10526D-03	0.45421D-01	0.446668D 00	0.446668D 00
0.19875D	02	0.0	0.23875D	02	0.10526D-03	0.45421D-01	0.446668D 00	0.446668D 00

THE APPROX. MARGINAL DISTRIBUTION OF X1

T	R•FREQ.	C•R•FREQ.	T	R•FREQ.	C•R•FREQ.
-0.20000D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.19375D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.18750D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.18125D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.17500D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.16875D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.16250D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.15625D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.15000D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.14375D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.13750D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.13125D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.12500D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.11875D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.11250D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.10625D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.10000D 01	0.0	0.0	0.0	0.98711D 00	0.98711D 00
-0.93750D 00	0.0	0.0	0.0	0.10526D -03	0.98737D 00
-0.87500D 00	0.0	0.0	0.0	0.10526D -03	0.98747D 00
-0.81250D 00	0.0	0.0	0.0	0.11875D 01	0.98768D 00
-0.75000D 00	0.0	0.0	0.0	0.12500D 01	0.98795D 00
-0.68750D 00	0.0	0.0	0.0	0.15789D -03	0.98811D 00
-0.62500D 00	0.0	0.0	0.0	0.16842D -03	0.98847D 00
-0.56250D 00	0.0	0.0	0.0	0.15625D 01	0.98916D 00
-0.50000D 00	0.0	0.0	0.0	0.14375D 01	0.98958D 00
-0.43750D 00	0.0	0.0	0.0	0.15263D -04	0.99042D 00
-0.25000D 00	0.0	0.0	0.0	0.16250D 01	0.63158D -03
-0.37500D 00	0.0	0.0	0.0	0.16875D 01	0.42105D -03
-0.31250D 00	0.0	0.0	0.0	0.15625D 01	0.84211D -03
-0.25000D 00	0.0	0.0	0.0	0.17500D 01	0.99105D 00
-0.18750D 00	0.0	0.0	0.0	0.18125D 01	0.99174D 00
-0.12500D 00	0.0	0.0	0.0	0.18750D 01	0.99237D 00
-0.62500D -01	0.0	0.0	0.0	0.19375D 01	0.10000D 01

THE APPROX. MARGINAL DISTRIBUTION OF X2

T	R•FREQ.	C•R•FREQ.	R•FREQ.	C•R•FREQ.
0.80000D 01	0.57895D-03	0.50142D 00	0.48000D-01	0.50142D 00
0.80625D 01	0.10526D-03	0.68421D-03	0.48632D-01	0.55005D 00
0.81250D 01	0.52632D-04	0.73684D-03	0.46632D-01	0.59668D 00
0.81875D 01	0.47368D-03	0.12105D-02	0.46632D-01	0.64332D 00
0.82500D 01	0.21053D-03	0.14211D-02	0.43789D-01	0.68711D 00
0.83125D 01	0.26316D-03	0.16842D-02	0.42526D-01	0.72963D 00
0.83750D 01	0.42105D-03	0.21053D-02	0.38579D-01	0.76821D 00
0.84375D 01	0.47368D-03	0.25789D-02	0.33895D-01	0.80211D 00
0.85000D 01	0.57895D-03	0.31579D-02	0.32842D-01	0.83495D 00
0.85625D 01	0.63158D-03	0.37895D-02	0.27947D-01	0.86289D 00
0.86250D 01	0.10526D-02	0.48421D-02	0.22053D-01	0.88495D 00
0.86875D 01	0.15789D-02	0.64211D-02	0.21368D-01	0.90632D 00
0.87500D 01	0.24211D-02	0.88421D-02	0.17842D-01	0.92416D 00
0.88125D 01	0.28421D-02	0.11684D-01	0.16684D-01	0.94084D 00
0.88750D 01	0.34211D-02	0.15105D-01	0.12316D-01	0.95316D 00
0.89375D 01	0.50000D-02	0.20105D-01	0.10211D-01	0.96337D 00
0.90000D 01	0.60526D-02	0.26158D-01	0.73158D-02	0.97068D 00
0.90625D 01	0.82632D-02	0.34421D-01	0.71579D-02	0.97784D 00
0.92500D 01	0.15421D-01	0.44632D-01	0.65263D-02	0.98437D 00
0.93125D 01	0.17579D-01	0.58789D-01	0.38947D-02	0.98826D 00
0.93750D 01	0.20211D-01	0.74211D-01	0.34737D-02	0.99174D 00
0.94375D 01	0.25211D-01	0.91789D-01	0.17368D-02	0.99347D 00
0.95000D 01	0.28947D-01	0.11200D 00	0.13158D-02	0.99663D 00
0.95625D 01	0.30789D-01	0.11200D 00	0.11500D 02	0.99768D 00
0.96250D 01	0.36632D-01	0.13721D 00	0.11375D 02	0.99532D 00
0.96875D 01	0.40053D-01	0.16616D 00	0.11438D 02	0.99832D 00
0.97500D 01	0.43421D-01	0.19695D 00	0.11563D 02	0.99900D 00
0.98125D 01	0.46421D-01	0.23358D 00	0.11625D 02	0.99926D 00
0.98750D 01	0.45000D-01	0.36347D 00	0.11688D 02	0.99953D 00
0.99375D 01	0.44947D-01	0.45342D 00	0.11813D 02	0.99979D 00
		0.11938D 02	0.10526D-03	0.99989D 00
		0.10000D 01	0.10526D-03	0.10000D 01

THE APPROX. MARGINAL DISTRIBUTION OF X3

T	R.FREQ.	C.R.FREQ.	T	R.FREQ.	C.R.FREQ.
0.50000D 00	0.12947D-01	0.12947D-01	0.45000D 01	0.38316D-01	0.34542D 00
0.62500D 00	0.52632D-04	0.13000D-01	0.46250D 01	0.40316D-01	0.38574D 00
0.75000D 00	0.0	0.13000D-01	0.47500D 01	0.39579D-01	0.42532D 00
0.87500D 00	0.52632D-04	0.13053D-01	0.48750D 01	0.39316D-01	0.46463D 00
0.10000D 01	0.15789D-03	0.13211D-01	0.50000D 01	0.38632D-01	0.50326D 00
0.11250D 01	0.47368D-03	0.13684D-01	0.51250D 01	0.43053D-01	0.54632D 00
0.12500D 01	0.31579D-03	0.14000D-01	0.52500D 01	0.38789D-01	0.58511D 00
0.13750D 01	0.73684D-03	0.14737D-01	0.53750D 01	0.42526D-01	0.62763D 00
0.15000D 01	0.52632D-03	0.15263D-01	0.55000D 01	0.39368D-01	0.66700D 00
0.16250D 01	0.10000D-02	0.16263D-01	0.56250D 01	0.39211D-01	0.70621D 00
0.17500D 01	0.11053D-02	0.17368D-01	0.57500D 01	0.36737D-01	0.74295D 00
0.18750D 01	0.78947D-03	0.18158D-01	0.58750D 01	0.33526D-01	0.77647D 00
0.20000D 01	0.14211D-02	0.19579D-01	0.60000D 01	0.32947D-01	0.80942D 00
0.21250D 01	0.26842D-02	0.22263D-01	0.61250D 01	0.27316D-01	0.83674D 00
0.22500D 01	0.30526D-02	0.25316D-01	0.62500D 01	0.25211D-01	0.86195D 00
0.23750D 01	0.32632D-02	0.28579D-01	0.63750D 01	0.21263D-01	0.88321D 00
0.25000D 01	0.41579D-02	0.32737D-01	0.65000D 01	0.19526D-01	0.90274D 00
0.26250D 01	0.46842D-02	0.37421D-01	0.66250D 01	0.17105D-01	0.91984D 00
0.27500D 01	0.58421D-02	0.43263D-01	0.67500D 01	0.14000D-01	0.93384D 00
0.28750D 01	0.83158D-02	0.51579D-01	0.68750D 01	0.13316D-01	0.94716D 00
0.30000D 01	0.10316D-01	0.61895D-01	0.70000D 01	0.11421D-01	0.95858D 00
0.31250D 01	0.10947D-01	0.72842D-01	0.71250D 01	0.90000D-02	0.96758D 00
0.32500D 01	0.12947D-01	0.85789D-01	0.72500D 01	0.62105D-02	0.97379D 00
0.33750D 01	0.16842D-01	0.10263D 00	0.73750D 01	0.64737D-02	0.98026D 00
0.35000D 01	0.18684D-01	0.12132D 00	0.75000D 01	0.57368D-02	0.98600D 00
0.38750D 01	0.24158D-01	0.14084D 00	0.76250D 01	0.36842D-02	0.98968D 00
0.40000D 01	0.26842D-01	0.21311D 00	0.80000D 01	0.18421D-02	0.99668D 00
0.41250D 01	0.29789D-01	0.24289D 00	0.81250D 01	0.12105D-02	0.99789D 00
0.42500D 01	0.30895D-01	0.27379D 00	0.82500D 01	0.42105D-03	0.99832D 00
0.43750D 01	0.33316D-01	0.30711D 00	0.83750D 01	0.16842D-02	0.10000D 01

THE APPROX. MARGINAL DISTRIBUTION OF X₄

T	R.FREQ.	C.R.FREQ.	T	R.FREQ.	C.R.FREQ.
-0.20000D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.19375D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.18750D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.18125D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.17500D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.16875D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.16250D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.15625D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.15000D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.14375D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.13750D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.13125D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.12500D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.11875D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.11250D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.10625D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.10000D .01	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.93750D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.87500D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.81250D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.75000D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.68750D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.62500D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.56250D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.50000D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.43750D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.37500D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.31250D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.25000D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.18750D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.12500D .00	0.0	0.0	0.0	0.99995D 00	0.99995D 00
-0.62500D-.01	0.0	0.0	0.0	0.99995D 00	0.99995D 00

SUMMARY OF THE RANDOM NUMBERS

THE MEAN VALUE OF R(1) IS 0.300D 01

THE MEAN VALUE OF R(2) IS 0.100D 01

THE MEAN VALUE OF R(3) IS 0.100D 01

THE MEAN VALUE OF R(4) IS 0.100D 01

THE MEAN VALUE OF R(5) IS 0.150D 02

THE MEAN VALUE OF R(6) IS 0.100D 02

THE VARIANCE-COVARIANCE MATRIX IS

0.100D-01	-0.199D-04	-0.669D-04	-0.274D-04	-0.635D-03	-0.628D-04
-0.199D-04	0.498D-02	0.129D-03	-0.176D-04	0.442D-03	0.106D-03
-0.669D-04	0.129D-03	0.506D-01	-0.235D-04	-0.358D-03	0.844D-04
-0.274D-04	-0.176D-04	-0.235D-04	0.997D-03	0.244D-03	-0.672D-04
-0.635D-03	0.442D-03	-0.358D-03	0.244D-03	0.639D 00	0.359D-03
-0.628D-04	0.106D-03	0.844D-04	-0.672D-04	0.359D-03	0.158D 00

B29957